Economic evaluation of treatment for externalizing disorders in adolescents

Connecting mental health and economics

Saskia Schawo

Economic Evaluation of Treatment for Externalizing Disorders in Adolescents

Connecting mental health and economics

Saskia Schawo

Colofon

Cover	photo: Saskia Schawo; lay-out: Carolien Laro
Lay-out	Rutger van Aken | persoonlijkproefschrift.nl
Print	Ridderprint BV | www.ridderprint.nl

Economic Evaluation of Treatment for Externalizing Disorders in Adolescents

Connecting mental health and economics

Economische evaluatie van behandelingen voor externaliserende problematiek in adolescenten
De brug slaan tussen mentale gezondheid en economie

Proefschrift

ter verkrijging van de graad van doctor aan de
Erasmus Universiteit Rotterdam
op gezag van de
rector magnificus

Prof.dr. R.C.M.E. Engels

en volgens besluit van het College voor Promoties.
De openbare verdediging zal plaatsvinden op

vrijdag 20 december 2019 om 11.30 uur

door

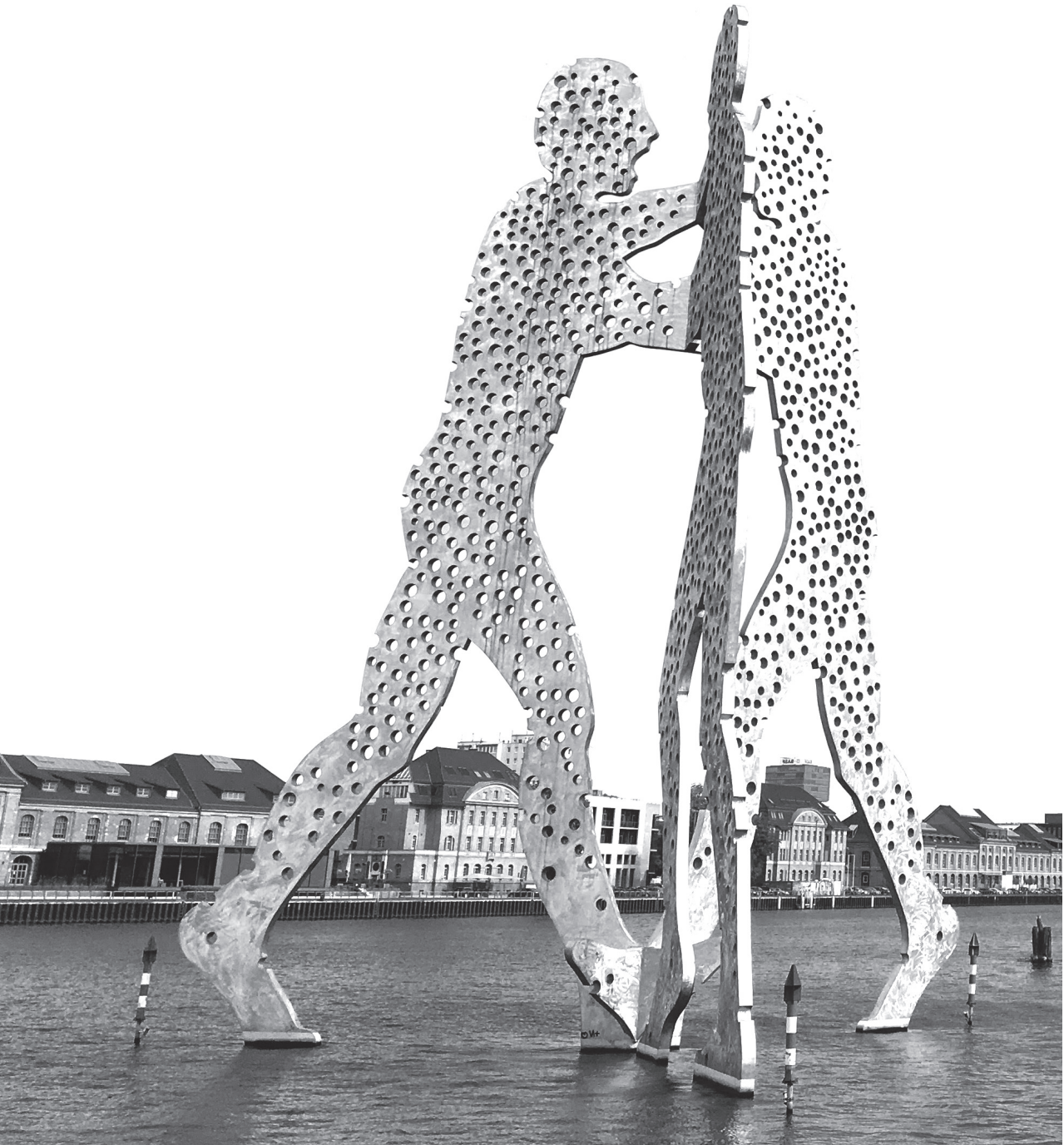
Saskia Jana Schawo
geboren te Bonn, Duitsland

Erasmus University Rotterdam

Promotiecommissie:

Promotor:	Prof.dr. W.B.F. Brouwer
Overige leden:	Prof.mr.dr. S.M.A.A. Evers Prof.dr. N.J.A. van Exel Dr. G.A. de Wit
Copromotor:	Dr. L. Hakkaart-van Roijen

Contents



Chapter 1

Introduction

Mental health

Mental disorders are common [1-3]. In the Netherlands, lifetime prevalence of one or more mental disorders in adults has been found to be 42.7% [2]. Being affected by a mental disorder can result in a high disease burden, both for the affected individuals as well as for their environment and society [4-7]. For the patient, this burden primarily relates to the quality of life losses related to mental disorders, which can be substantial [8-10]. For the patients' environment, mental health problems may lead to disturbed personal relationships, stress, and strain on caregivers [11, 12]. The societal burden related to mental disorders includes aspects like the high health care costs of treating mental disorders [13,14, 6], losses of productivity due to absenteeism and reduced working capacity [15, 16], as well as pressure on sectors such as social care, education and criminal justice [17-19]. Due to these broad impacts of mental disorders at several levels and in different sectors, specifying the overall burden of the disorders is a complex and challenging task.

Externalizing mental disorders in adolescents

The multifaceted and substantial impact of mental disorders also holds for externalizing mental disorders in adolescents. Externalizing disorders are mental disorders, like Attention-Deficit-Hyperactivity Disorder (ADHD), Conduct Disorder (CD) and Antisocial Personality Disorder, which are outwardly directed and therefore, by definition, do not only affect the patient, but also his or her environment. These disorders are therefore associated with a particularly wide variety of costs and effects. Symptoms of these disorders range from concentration problems and restlessness or disobedience to aggressive behavior, violence and substance use [20-22]. Patients may experience significant impairments in social, academic or occupational functioning [20, 23-25]. This can create high individual and societal burdens, both within and beyond the health care sector. Especially in younger patients, this burden will include aspects such as school performance, criminal activity, disturbed relationships with parents and siblings, and so forth. Several psychotherapeutic and psychosocial interventions for treatment of externalizing disorders exist, including, for example, Functional Family Therapy (FFT), Multisystemic Therapy (MST), Multidimensional Family Therapy (MDFT), Cognitive Behavioral Therapy (CGT), parent training, and community-based interventions [26-29]. Some of these interventions are specifically directed at improving patients' interactions with the various systems around them (i.e., parents, siblings, peers, teachers, colleagues, neighbors, 'society as a whole').

Budget limitations

The broad and diverse impacts of externalizing disorders in adolescents and related intervention strategies pose clear challenges for evaluating the (costs and) effects of such interventions. Nonetheless, such evaluations are needed. To date, limited information is available regarding the costs and effects of these interventions. Yet, this information is of high importance for medical and policy decisions concerning preferred treatment and the funding thereof. In recent years, changes in government policy, technological advances, increasing wealth and population ageing have contributed to increasing health care expenditures [30-32]. In the Netherlands, also as a result of changes in government policy, mental health care expenses rose more (+105%) than overall health care expenditures (+49%) [33] between 2000 and 2010. This raises questions regarding the sustainability of such expenditures and growth rates as well as the justification of how budgets are spent, given that resources are limited and more spending on (mental) health care has opportunity costs inside and outside the health care sector. Ideally, limited resources would be allocated in the most efficient way, so that they optimally contribute to improving overall health or welfare. However, a lack of information on costs and effects of interventions makes it difficult to provide evidence-based advice to policymakers on which interventions contribute optimally to their goals.

Health economic evaluation has become a commonly used tool to inform such policy and budget decisions [34]. Yet, to what extent the classical health economic methodology sufficiently and adequately captures the broad costs and effects of mental health interventions in adolescents with externalizing disorders remains a matter of debate. This has been previously highlighted for complex mental health conditions and mental disorders in general by Brazier et al. [35, 36] and Knapp et al. [37]. The lack of information on costs and effects of mental interventions, together with questions concerning the suitability of the common methodology used in economic evaluations to assess these, is at odds with the societal and scientific relevance of providing more insight into these issues. This is especially the case when policy makers wish to stimulate effective and cost-effective treatments of adolescents with mental disorders (and externalizing disorders in particular).

Economic evaluations

In health economic analysis costs are compared to the benefits of an intervention. The aim of the analysis, when taking a societal perspective, is to answer the

question as to whether the benefits exceed the costs, thus demonstrating that intervention results in an increase in societal welfare. Several types of health economic evaluation exist. Classical cost-benefit analysis (CBA) compares costs and benefits both expressed in monetary units. It directly answers the question whether benefits exceed the costs of the intervention. Whereas CBA is more common in other sectors, it is used less often in health care. There, cost-effectiveness analysis (CEA) is more common. CEA compares costs expressed in monetary units with effects expressed in a unit relevant to the outcome of the intervention (i.e., costs per life year gained, costs per hip fracture avoided, costs per point decrease on some clinical scale, etc.). The advantage of doing this is that such outcomes, which cannot easily be expressed in monetary terms, relate well to the clinical practice and can still be evaluated. Yet, a disadvantage is that CEA uses diverse outcomes across settings, which limits comparability of results and hence consistency of decision making. CEA can be quite useful in a clinical setting therefore, but when aiming to inform societal decision-making, comparability between different interventions is important in order to judge which intervention contributes most to health and welfare in relation to its costs. Cost-utility analysis (CUA) is more suitable for this goal. In CUAs, costs are measured in terms of monetary units and outcomes are evaluated in terms of quality adjusted life years (QALYs). The QALY measure comprises both length and quality of life, the latter expressed in quality of life weights typically based on preferences in the general public for different health states. These health states are measured using (generic or disease-specific) health-related quality of life measures like the EQ-5D [38]. Using the QALY, outcomes of analyses become comparable across interventions without (directly) monetarizing these effects. In the Netherlands, like in several other countries (e.g., Canada, Australia, and the UK), specific guidelines for the performance of economic evaluations have been developed. These guidelines suggest CUA as the preferred methodology [39, 40].[1]

Connecting economic evaluations and mental health

Performing health economic evaluations in mental health care, regardless of whether they take the form of a CBA, CEA or CUA, is challenging. Measuring outcomes of mental health interventions is not yet as common as the assessment of physical symptoms in medical care [41]. It is often difficult and even contentious to measure and value the broad benefits of mental health interventions. The

1 The terms CEA and CUA are often used interchangeably. CEA is generally used when natural units are involved and CUA when outcomes are measured in terms of QALYs. In this dissertation, we use the expression CEA as an overarching term unless otherwise indicated in the text.

measurement and valuation of these broad benefits is hampered by the fact that standardized instruments specifically designed for this purpose and suitable for inclusion in economic evaluations are lacking, and that there are questions regarding whether conventional QALY measures are adequate in this context [35-37]. Similar concerns exist regarding identifying, measuring and valuing the broad variety of societal costs (and savings) associated with mental health disorders and their treatments. It has been previously suggested that the current methodology of CUA mainly focuses on improvements in health and may insufficiently capture broader effects of interventions [42]. This criticism seems relevant for sectors like social care and elderly care, but also for mental health, with broad personal and societal impacts, rendering traditional outcome measures like QALY instruments and included cost-categories potentially insufficient for a full welfare economic assessment of these interventions. Given these concerns, there is an ongoing debate about the suitability of the current methodology of economic evaluations, also in the context of mental health [35, 43]. The discussion is wide-ranging and concerns issues such as the QALY not being able to capture treatment goals more broadly than the health dimension alone [37, 44], the inclusion of effects related to work or family functioning [35], and the inclusion of broader societal costs and benefits such as productivity losses and informal care [45, 46].

Performing economic evaluations of externalizing disorders in adolescents

CUAs of interventions to treat externalizing disorders are scarce [47]. In line with what was mentioned above, measuring the effects of interventions for externalizing disorders in adolescents can be considered particularly challenging, as (intended) treatment effects may be broader than health gains alone. Due to the interactional characteristics of both the disorders and the interventions, interventions may (intend to) affect the system around the patient as well (e.g., improvements in social interactions, school performance, reduction of violence or substance use, etc). Such broader impacts may result in changes in the health or wellbeing of patients and their families, as well as in societal costs or savings. Therefore, particular attention would be required to capture these broad costs and effects when determining the cost-effectiveness of an intervention. Furthermore, long-term effects play an important role in this patient population as treatment during adolescence may prevent problems later on in life, such as delinquency or the need for more intensive and complex treatments (for the patient, the system or victims). These issues need particular attention when evaluating the costs and

effects of interventions for externalizing disorders in adolescents. Within the current health economic methodology this long-term horizon can be taken into consideration using health economic modeling techniques, which may however be complex in light of the above-mentioned contextual costs and effects of mental disorders in adolescents with externalizing disorders.

Objective

The overall aim of this dissertation is therefore to explore different ways of improving the methodology of economic evaluations of interventions for externalizing disorders in adolescents. This thesis takes first (explorative) steps in addressing this issue and bridging the gap between the specific goals of interventions for adolescents with externalizing behavioral disorders and conventional health economic methodology. We investigate this by first applying conventional methodology (using health economic modeling techniques, value of information analysis and the QALY as outcome measure), then using a simple one-dimensional alternative outcome measure, and finally developing a broader, preference based outcome measure. Ultimately, this thesis aims to contribute to the improvement of health economic evaluations of interventions targeted at externalizing mental disorders, making such evaluations more valuable for policymaking.

In this thesis, a number of steps will be taken in designing a comprehensive outcome measure, potentially useful in evaluations of interventions aimed at treating externalizing disorders. We note upfront that having a separate measure for this context necessarily compromises comparability of results of economic evaluations across different settings. However, it also improves the comprehensiveness of the captured benefits deemed important in the context of mental health. Hence, in this search, we may sacrifice part of the comparability between interventions in exchange for a more comprehensive and meaningful outcome measure.

Outline

This thesis consists of different chapters, which are all based on independently readable papers. Each of the chapters addresses a specific research question, related to the overall aim of this thesis. These research questions are listed below.

Chapter 2: *How can a cost-effectiveness analysis for pharmacological treatment of an externalizing disorder (ADHD) be performed, including*

	consideration of relevant broader societal impacts while using conventional health economic methodology?
Chapter 3:	*What do we know about the cost-effectiveness of systemic interventions for delinquency and substance use?*
Chapter 4:	*Can we perform a cost-effectiveness analysis of a systemic intervention for delinquency in adolescents using Criminal Activity Free Years as outcome measure?*
Chapter 5:	*Can we perform a Value of Information analysis based on the cost-effectiveness analysis using Criminal Activity Free Years as outcome measure, to inform future research?*
Chapter 6:	*Which treatment effects should be captured in economic evaluations of systemic interventions in adolescents according to clinicians and do existing QALY measures capture these?*
Chapter 7:	*Which outcome measures are currently used to measure the effects of systemic interventions in clinical research and could these be used in cost-utility analyses?*
Chapter 8:	*Is it possible to obtain societal preference-weights for a comprehensive multidimensional outcome measure to be used in economic evaluations of systemic interventions targeted at adolescents with problems of substance use and delinquency?*

The thesis outline is as follows. Chapter 1 provided a background on economic evaluations in relation to specific characteristics and challenges of externalizing disorders and interventions aimed at these. It also introduced the goal of this dissertation.

In chapters 2 and 3 the 'standard approach' of economic evaluation in health care is applied in the context of interventions in the field of mental health, and the literature is reviewed to learn more about the outcomes of such applications for interventions for externalizing behavioral disorders. Specifically, chapter 2 reports the results of a classical probabilistic CUA, investigating treatment of children and adolescents with ADHD with short-acting or long-acting methylphenidate. The analysis applies commonly used health-related outcomes, but includes some relevant broader societal aspects. Chapter 3 provides an overview of what is known regarding cost-effectiveness of interventions for externalizing behavioral disorders, based on a systematic literature review.

In chapters 4 and 5, we highlight how economic evaluations of mental health interventions could be performed using a tailored, yet very simple, outcome measure: Criminal Activity Free Years (CAFY). In chapter 4, a classical probabilistic CEA model is used to evaluate Functional Family Therapy (FFT), a systemic intervention, compared to treatment as usual, using the CAFY. Chapter 5 builds on the results from chapter 4 by investigating the value of future research on specific parameters of this classical CEA model, using value of information analysis. Obviously, while showing that standard methodology can be applied using a context specific outcome measure, the measure used in these two chapters is crude, narrow and (too) simple. For instance, it only considers criminal activity as a relevant outcome and assigns the same weight to different delinquent activities (i.e., the same weight for stealing a bike as for murder).

Given these limitations, chapters 6, 7 and 8 further investigate how the existing CUA methodology could be improved in the context of mental health interventions. Ideally, a comprehensive multidimensional outcome measure with societal preference-weights would exist that could be used in this context. Chapter 6 first examines, based on interviews with clinicians, which effects according to these professionals should be captured in cost-effectiveness analyses of systemic interventions, also given the envisioned therapeutic goals, and whether current generic QALY measures capture these. Chapter 7 summarizes the results of a systematic literature review of outcome measures used in evaluations of systemic interventions. We also investigate whether one of the existing measures found in the review captures all relevant outcomes of systemic interventions and can be considered suitable for use in CUA. Chapter 8 determines societal preference-weights for a (shortened) multidimensional instrument that was labeled as being promising in chapter 7, as to make it suitable for use in CUA of systemic interventions.

Chapter 9 presents the main conclusions of this thesis, reflects on the results and provides suggestions for further research.

References

[1] Bijl RV, Ravelli A, Van Zessen G. Prevalence of psychiatric disorder in the general population: results of The Netherlands Mental Health Survey and Incidence Study (NEMESIS). *Social psychiatry and psychiatric epidemiology*. 1998;*33*(12):587-595.

[2] de Graaf R, Ten Have M, van Gool C, van Dorsselaer S. Prevalence of mental disorders and trends from 1996 to 2009. Results from the Netherlands Mental Health Survey and Incidence Study-2. *Social psychiatry and psychiatric epidemiology*. 2012;*47*(2):203-213.

[3] Singleton N, Bumpstead R, O'Brien M, Lee A, Meltzer H. *Psychiatric Morbidity Among Adults Living in Private Households, 2000: Summary Report*. London: Office for National Statistics; 2000.

[4] Simon GE, Ornel J, Von Korff M, Barlow W. Health care costs associated with anxiety and depressive disorders in primary care. *American Journal of Psychiatry*. 1995;*152*:353-357.

[5] Slobbe LC, Smit JM, Groen J, Poos MJ, Kommer GJ. *Kosten van ziekten in Nederland 2007-2010*, Rijksinstituut voor Volksgezondheid en Milieu: Bilthoven; 2011.

[6] Wykes T, Haro JM, Belli SR, Obradors-Tarragó C, Arango C, Ayuso-Mateos JL et al. Mental health research priorities for Europe. *The Lancet Psychiatry*. 2015;*2*(11):1036-1042.

[7] Kessler R, Barber C, Birnbaum H, Frank R, Greenberg P, Rose R et al. Depression in the workplace: effects on short-term disability. *Health Affairs. 1999;18*:163-171.

[8] Rapaport MH, Clary C, Fayyad R, Endicott J. Quality-of-life impairment in depressive and anxiety disorders. *American Journal of Psychiatry*. 2005;*162*(6):1171-1178.

[9] Spitzer R, Kroenke K, Linzer M, Hahn S, Williams J, Verloin de Gruy F et al. Health-related quality of life in primary care patients with mental disorders. Results from the PRIME ^ MD 1000 Study. *JAMA. 1995;274*:1511-1517.

[10] Wittchen HU, Carter RM, Pfister H, Montgomery SA, Kessler RC. Disabilities and quality of life in pure and comorbid generalized anxiety disorder and major depression in a national survey. *International clinical psychopharmacology*. 2000;*15*(6):319-328.

[11] Al Janabi H, Frew E, Brouwer W, Rappange D, Van Exel J. The inclusion of positive aspects of caring in the Caregiver Strain Index: tests of feasibility and validity. *International journal of nursing studies*. 2010;*47*(8):984-993.

[12] Hoefman R, Payakachat N, van Exel J, Kuhlthau K, Kovacs E, Pyne J, Tilford JM. Caring for a child with autism spectrum disorder and parents' quality of life: application of the CarerQol. *Journal of autism and developmental disorders*. 2014;*44*(8):1933-1945.

[13] Hakkaart-van Roijen L, Zwirs BWC, Bouwmans C, Tan SS, Schulpen TWJ, Vlasveld L et al. Societal costs and quality of life of children suffering from attention deficient hyperactivity disorder (ADHD). *European child & adolescent psychiatry*. 2007;*16*(5):316-326.

[14] Soeteman DI, Roijen LHV, Verheul R, Busschbach JJ. The economic burden of personality disorders in mental health care. *Journal of Clinical Psychiatry*. 2008;*69*(2):259.

[15] Hakkaart-van Roijen L, Hoeijenbos MB, Regeer EJ, Ten Have M, Nolen WA, Veraart CPWM, et al. The societal costs and quality of life of patients suffering from bipolar disorder in the Netherlands. *Acta Psychiatrica Scandinavica*. 2004;*110*(5):383-392.

[16] van der Feltz-Cornelis CM, Hoedeman R, de Jong FJ, Meeuwissen JA, Drewes HW, van der Laan NC et al. Faster return to work after psychiatric consultation for sicklisted employees with common mental disorders compared to care as usual. A randomized clinical trial. *Neuropsychiatric disease and treatment*. 2010;*6*:375.

[17] Barrett B, Byford S, Seivewright H, Cooper S, Duggan C, Tyrer P. The assessment of dangerous and severe personality disorder: service use, cost, and consequences. *The Journal of Forensic Psychiatry & Psychology*. 2009;*20*(1):120-131.

[18] Cary M, Butler S, Baruch G, Hickey N, Byford S. Economic evaluation of multisystemic therapy for young people at risk for continuing criminal activity in the UK. *PloS one*. 2013;*8*(4):e61070.

[19] De Socio J, Hootman J. Children's mental health and school success. *The Journal of School Nursing*. 2004;*20*(4):189-196.

[20] American Psychiatric Association. Diagnostic and statistical manual of mental disorders (DSM-IV-TR). *American Psychiatric Association*.; 2000.

[21] King SM, Iacono WG, McGue M. Childhood externalizing and internalizing psychopathology in the prediction of early substance use. *Addiction*. 2004;*99*(12):1548-1559.

[22] Krueger RF, Markon KE, Patrick CJ, Iacono WG. Externalizing psychopathology in adulthood: A dimensional-spectrum conceptualization and its implications for DSM-V. *Journal of abnormal psychology*. 2005;*114*(4):537.

[23] Bongers IL, Koot HM, van der Ende J, Verhulst FC. Predicting young adult social functioning from developmental trajectories of externalizing behaviour. *Psychological Medicine*. 2008;*38*(7):989-999.

[24] Knapp M, King D, Healey A, Thomas C. Economic outcomes in adulthood and their associations with antisocial conduct, attention deficit and anxiety problems in childhood. *Journal of Mental Health Policy and Economics*. 2011;*14*(3):137-147.

[25] Colman I, Murray J, Abbott RA, Maughan B, Kuh D, Croudace TJ et al. Outcomes of conduct problems in adolescence: 40 year follow-up of national cohort. *BMJ. 2009;338*:a2981.

[26] Farmer EM, Compton SN, Burns JB, Robertson E. Review of the evidence base for treatment of childhood psychopathology: Externalizing disorders. *Journal of consulting and clinical psychology*. 2002;*70*(6):1267.

[27] Alexander JF, Robbins MS. Functional family therapy. In *Clinical handbook of assessing and treating conduct problems in youth* (pp. 245-271). Springer, New York, NY; 2011.

[28] Henggeler SW. Multisystemic therapy: An overview of clinical procedures, outcomes, and policy implications. *Child Psychology and Psychiatry Review*. 1999;*4*(1):2-10.

[29] Liddle HA, Dakof GA, Parker K, Diamond GS, Barrett K, Tejeda M. Multidimensional family therapy for adolescent drug abuse: Results of a randomized clinical trial. *The American journal of drug and alcohol abuse*. 2001;*27*(4):651-688.

[30] Christiansen T, Bech M, Lauridsen J, Nielsen P. Demographic changes and aggregate health-care expenditure in Europe. CEPS; 2006.

[31] Weisbrod BA. The health care quadrilemma: an essay on technological change, insurance, quality of care, and cost containment. *Journal of economic literature*. 1991;29(2):523-552.

[32] Reinhardt UE. Does the aging of the population really drive the demand for health care?. *Health Affairs*. 2003;*22*(6):27-39.

[33] Statistics Netherlands (CBS). Database of the CBS. Available: https://opendata.cbs.nl/statline/#/CBS/nl/navigatieScherm/thema. Accessed 2019.

[34] Drummond, MF. *Methods for the economic evaluation of health care programmes*. Oxford university press; 2005.

[35] Brazier J. Is the EQ–5D fit for purpose in mental health?. The Britisch Journal of Psychiatry. 2010;*197*(5):348-349

[36] Brazier J, Tsuchiya A. Improving cross-sector comparisons: going beyond the health-related QALY. *Applied health economics and health policy*. 2015;*13*(6):557-565.

[37] Knapp M, Mangalore R. The trouble with QALYs…. *Epidemiology and Psychiatric Sciences*. 2007;*16*(4):289-293.

[38] Brooks R, Rabin R, De Charro F. (Eds.). *The measurement and valuation of health status using EQ-5D: a European perspective: evidence from the EuroQol BIOMED Research Programme*. Springer Science & Business Media; 2013.

[39] Zorginstituut Nederland. *Guideline for Economic Evaluations in Health Care.* Diemen: Zorginstituut Nederland; 2016.

[40] National Institute for Health and Care Excellence. *Guide to the methods of technology appraisal 2013*. London; 2013.

[41] Rose M, Devine J. Assessment of patient-reported symptoms of anxiety. *Dialogues in clinical neuroscience*. 2014;*16*(2):197.

[42] Coast J. Is economic evaluation in touch with society's health values?. *BMJ: British medical journal*. 2004;*329*(7476):1233-1236.

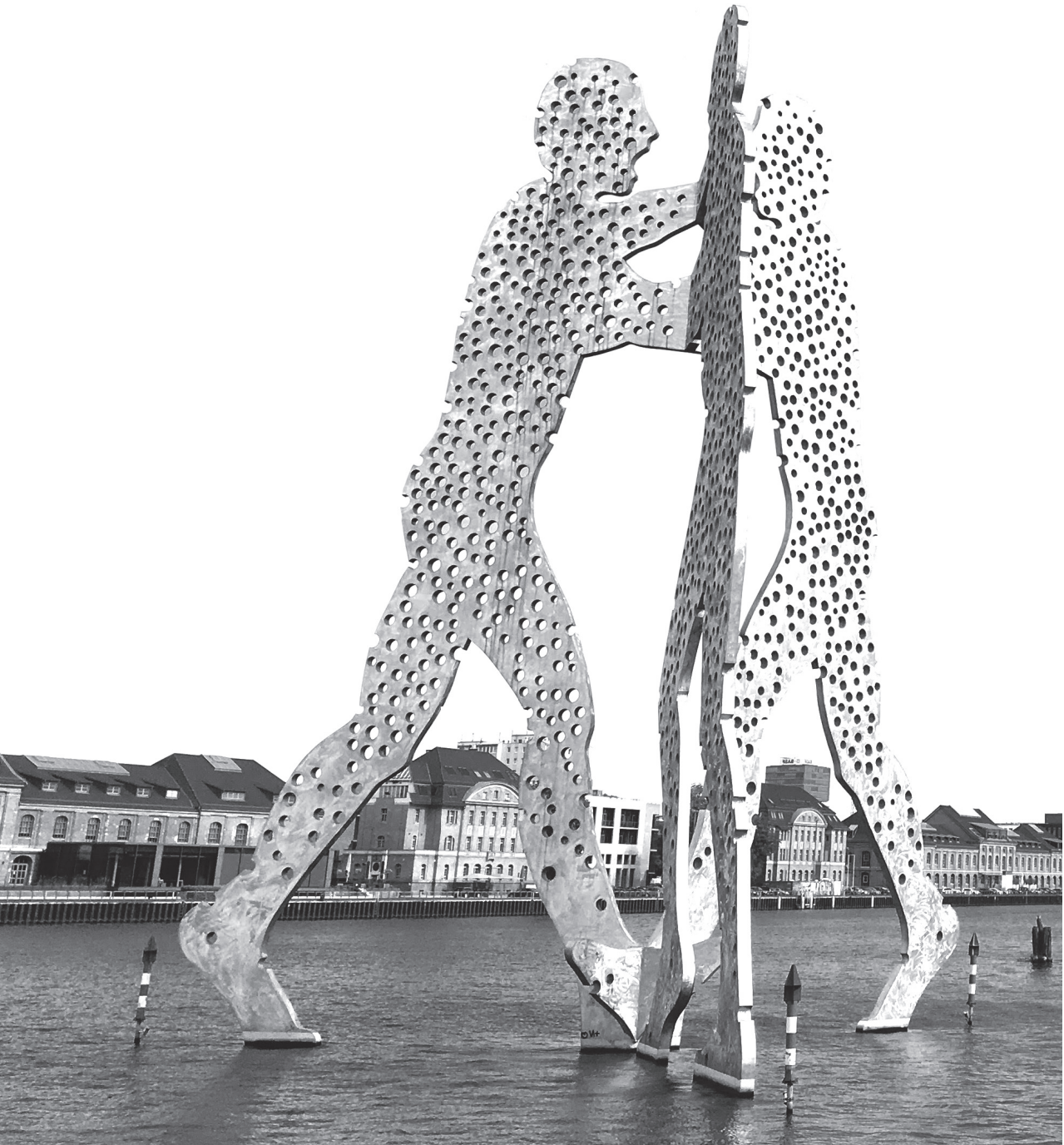
[43] Ryan M, Gerard K. Inclusiveness in the health economic evaluation space. *Social Science & Medicine; 2014;108*:248-251.

[44] Al-Janabi H, Flynn TN, Coast J. Development of a self-report measure of capability wellbeing for adults: the ICECAP-A. *Quality of Life Research*. 2012;*21*(1):167-176.

[45] Al-Janabi H, Flynn TN, Coast J. Estimation of a preference-based carer experience scale. *Medical Decision Making*. 2011;*31*(3):458-468.

[46] Pyne JM, French M, McCollister K, Tripathi S, Rapp R, Booth B. Preference-weighted health-related quality of life measures and substance use disorder severity. *Addiction*. 2008;*103*(8):1320-1329.

[47] Goorden M, Schawo SJ, Bouwmans-Frijters CA, van der Schee E, Hendriks VM, Hakkaart-van Roijen L. The cost-effectiveness of family/family-based therapy for treatment of externalizing disorders, substance use disorders and delinquency: a systematic review. *BMC psychiatry*. 2016;*16*(1):237.



Chapter 2

Probabilistic Markov Model Estimating Cost Effectiveness of Methylphenidate Osmotic-Release Oral System Versus Immediate-Release Methylphenidate in Children and Adolescents: Which Information is Needed?

Based on Schawo, S., van der Kolk, A., Bouwmans, C., Annemans, L., Postma, M., Buitelaar, J., van Agthoven, M. & Hakkaart-van Roijen, L.

PharmacoEconomics. 2015; 33(5): 489-509.

Introduction

An increasing incidence of attention deficit hyperactivity disorder (ADHD) in children [1] and high use of pharmacological treatments [2] have become relevant issues for policymakers and mental health professionals. It is yet unclear whether the increase in incidence is due to changes in true numbers of patients or whether numbers appear higher as a consequence of differences in diagnosis or recall of parents [3]. The high number of young ADHD patients results in significant societal costs [4]. Evidence from literature suggests that 50-70% of those suffering from ADHD in childhood also experience ADHD as an adult [5, 6]. Hence costs are not limited to the short term; ADHD may also result in lower household income, mental and physical dysfunction, comorbidities and increased health consumption later on in life [6, 7] as well as increased health care consumption and productivity losses of household members [8].

First-choice medication for treatment of ADHD in the Netherlands is the stimulant methylphenidate (MPH) [9]. MPH is available as a short-acting as well as a more costly long-acting formulation. Different formulations are available from a wide selection of brands and in different strengths. Short-acting MPH requires accurate medication intake 2-5 times a day [9]. Consequently, medication intake may require high effort and impose practical difficulties, for example, on children attending school. The long-acting formula has been developed to overcome those practical problems of medication intake and compliance by using a once-a-day treatment scheme [10]. Existing clinical studies suggest no significant difference between the efficacy of short-acting and long-acting MPH under the assumption of full therapy compliance [10-12]. However, it has been shown that lower frequency of medication intake is correlated with better treatment compliance [13]. Long-acting MPH has shown to be associated with better treatment continuity [14, 15]. Kemner and Lage [14] found patients treated with long-acting MPH to be subject to less breaks in medication use, fewer medication switches and a longer period on intended therapy. Marcus et al. [15] stated that the treatment duration of patients with long-acting MPH was on average longer than for patients treated with short-acting MPH. Long-acting formulations of MPH have also been proven to result in superior compliance in patients when compared to the short-acting formulation [16-18], hence, possibly leading to better effectiveness than the short-acting formulation.

However, it is not evident whether the effect of long-acting formulations of MPH can justify the higher costs. Given the scarce financial resources in health

care, cost-effectiveness analyses have become essential to inform policymakers' choices between competing treatments and to provide founded recommendations to clinicians within clinical guidelines. However, evidence in the form of recent state-of-the-art health economic evaluations of ADHD treatment in children is limited. Furthermore, there is increasing debate on whether it is sufficient to purely evaluate interventions on the basis of costs and effects in the domain of health care and limit these to the patient alone [19]. Authors of recent publications emphasized the lack of economic studies on ADHD in children and adolescents with a broader societal perspective [20-22]. Bernfort et al. [21] found that most often societal costs were not included in economic evaluations of ADHD. Wu et al. [22] performed a systematic literature review on health care costs of family members of children with ADHD and found those costs to be higher than those of families without a child with ADHD. Beecham [20] stated that "economic evaluation of interventions for child and adolescent psychiatric disorders has lagged some way behind its adult counterpart." She expressed the need for a broader perspective as to reflect the various effects of psychiatric disorders in children and adolescents [20]. Evidence from the literature on meningitis [23] suggests that 'spillover' health effects on family may constitute as much as 48% of the health effects on the patient. As ADHD can be considered especially stressful on the direct environment of the patient, such as parents, siblings, friends or schoolmates, this percentage may even be larger for patients with ADHD. Hence, the inclusion of broader societal effects and costs is considered necessary [22].

Bernfort et al. [21] recommended the use of a health economic Markov model to determine the long-term costs and effects of ADHD. However, the authors stated that sufficiently detailed data (especially on long-term consequences of ADHD) was scarce or unavailable [21]. King et al. [24] expressed their concerns on the limited availability of effectiveness estimates and utility values, possibly due to scarcity of clinical data. Among the health economic evaluations that have been performed to evaluate various pharmacological treatments of ADHD are analyses based on decision analytic trees [25] and cost-of-illness calculations [26]. A small number of evaluations have been performed based on more advanced health economic (Markov) models [24, 27, 28]. However, there is a lack of more recent studies in the field. An economic evaluation on long-acting MPH osmotic release systems (OROS) versus short-acting MPH immediate release (IR) suggested better cost-effectiveness of OROS (hereafter referred to as the Faber model) [29]. However, that evaluation was limited compared with the current standard of HE modeling as a deterministic model was employed and

only effects limited to the patient were included. Hence, clear health economic recommendations on the cost-effectiveness of OROS compared to IR based on a broad societal perspective are still lacking.

Knowledge of the cost effectiveness of treatment options for children with ADHD is essential in order to inform policymakers and enable the formulation of specific recommendations in clinical guidelines. In the case of MPH, it would be desirable to provide clear recommendations on which formulation is to be preferred under which circumstances, founded on sound and comprehensive health economic evidence. This study aims to contribute to this goal. We perform a cost-effectiveness evaluation of OROS versus IR in line with current health economic methodology, based on the Faber model [29], but with a probabilistic model update, enhanced model structure, updated input parameters (including utility values) and a broader societal perspective (i.e. we considered criminal justice costs, educational costs, employment disadvantages, out-of-pocket-expenses, medical and productivity costs and utility values of the caregiver). Additionally, we provide specific recommendations for future data collection, which would be valuable to further increase the validity of the model outcomes.

Methods

We evaluated the cost-effectiveness of OROS compared with IR for patients with suboptimal response to IR. The structure of the probabilistic Markov model and its parameters were defined according to the Dutch guidelines for pharmacoeconomic evaluation [30].

According to health economic standards, a societal perspective was taken to reflect costs and effects on patients, their parents and society as a whole [31]. We searched literature on a broad range of cost categories for relevance and feasibility of inclusion in the model (i.e. criminal justice costs, lower income, out-of pocket expenses of the patient as well as health care costs and productivity costs of caregivers). Direct medical and non-medical costs as well as spillover effects on caregivers were included in the model.

Consultation of experts

As part of this study, a panel of experienced psychiatrics from various regions in the Netherlands was consulted (table 1). These experts were asked to provide feedback on the model structure, input and model assumptions as well as estimates of transition probabilities. Transition probabilities were retrieved

in accordance with the Delphi panel requirements [30], and other issues were discussed individually. After discussion with the expert panel on, among others, the definition of health states and the cycle length of the probabilistic model, the cycle length was chosen to remain unchanged and the model states were slightly adapted as opposed to the Faber model [29] to better match patient characteristics, illness and treatment approach.

Table 1 | Consulted experts

Expert	Gender	Age (years)	Specialism	Sub specialism	Years experience in mental health	Average number of patients with ADHD from 6 to 18 yrs seen per month	Years experience with ADHD medication	Average number of patients seen/ month
1	M	55	Child- and youth psychiatrist	None	24	90	16	105
2	M	52	Child- and youth psychiatrist	Hospital, child psychiatry and ADHD	22	>30	16	>100
3	F	43	Child- and youth psychiatrist	ADHD/ODD/ticks	13	45	10	50
4	M	55	Child psychiatrist	Neuropsychiatry	29	50	22	80

ADHD attention deficit hyperactivity disorder, ODD oppositional defiant disorder

General model characteristics

The probabilistic model was based on the existing deterministic model by Faber et al. [29]. Model type, model state definitions, time horizon, model parameters and model input (including utility values) were updated to enhance the existing model and to comply with current health economic methodology.

Table 2 | Current model vs. Faber model

Current model	**Faber model** [29]
General model type: Markov model *Perspective*: societal *Cycle length:* 1 day *Resource use estimates:* expert panel (Faber et al. [29]) *Outcomes:* expressed as cost/QALY	
Utility estimates: Patient and caregiver	*Utility estimates:* Patient
Reference of utility estimates: van der Kolk et al. [57]	*Reference of utility estimates:* Secnik et al. [32]
Specific model type: probabilistic	*Specific model type:* deterministic
Model states OROS/IR: optimal, suboptimal, treatment stopped, remission	*Model states OROS:* optimal, non compliance, treatment stopped, functional remission *Model states IR:* optimal, suboptimal, treatment stopped, functional remission
Patient age when entering model: 6 years	*Patient age when entering model:* 8 years
Time horizon: 12 years	*Time horizon:* 10 years
Transition rate estimates: Delphi panel of experts	*Transition rate estimates:* various sources (literature and expert opinion)
Cost categories: Patient: Medication costs, consultation costs, intervention costs, special education costs Caregiver: Medical costs, production losses	*Cost categories:* Patient: Medication costs, consultation costs, intervention costs, special education costs
Cost parameter values: 2014 EUR	*Cost parameter values:* 2005 EUR

As the Faber model was limited to a deterministic decision-analytic model with sensitivity analyses, we chose a more advanced probabilistic approach. The consideration of uncertainty increasingly gains importance, as shown in several guidelines, of which one explicitly suggests the use of probabilistic sensitivity

analysis [33]. Therefore, input parameters were set to vary according to prior distributions as to introduce parameter uncertainty in the model.

Furthermore, we refined and improved model state definitions. Where Faber et al. [29] considered five model states (optimal, suboptimal, treatment stopped, functional remission and non-compliance), with different states applicable for different treatment conditions, the current model defined four model states (optimal, suboptimal, treatment stopped and remission) consistent across treatment conditions.

The time horizon of the model was slightly adjusted in the current model. Patients entered the Faber model [29] at 8 years of age and remained in the model for 10 years. In the current model, we redefined the starting age of patients entering the model to 6 years and extended the time horizon to 12 years in order to be in line with the treatment guidelines for ADHD [9]. The relevant patient population was defined as patients within this age group who initially had responded suboptimally to IR because of incorrect intake of medication (i.e. missing doses of medication due to administrative burden). To simulate a randomized population, it was assumed that half of the initial patient population continued to receive IR and the other half switched to treatment with OROS when entering the model.

Within the current model, the assumed cycle length was one day and was consistent with the set-up of the Faber model. The panel of experts (table 1) indicated that a cycle length in line with the prescription regimen of a day would be most appropriate and consistent as non-compliance to medication would, on average, result in a change in behavior on the same day for almost all children, with only few exceptions. This cycle length implies that an improvement or worsening of compliance can occur on a daily basis and symptoms and costs change accordingly after one day. In reality, costs may adjust less quickly than effects, resulting in less volatility in costs than assumed in the model.

The prescribed dosage of medication was assumed optimal for all patients based on age and metabolism. In line with the Multimodel Treatment of attention deficit hyperactivity disorder (MTA) study [34] and expert comments, a mean of three doses IR per day and one dose OROS per day were assumed.

Costs and effects were discounted at 4% and 1.5% respectively, according to the Dutch guidelines for pharmacoeconomic research [30].

Model states

The effect of medication was evaluated in terms of ADHD symptoms and behavioral change. The model distinguished four different health states (table 3). The definition of the health states was based on the Faber model [29] and enhanced with feedback from the expert panel. Where Faber et al. [29] made a distinction between a suboptimal state for treatment with IR and the state of non-compliance for treatment with OROS, the updated model made use of a consistent health state definition over treatments. The non-compliance state was replaced by the suboptimal state, now defined as a state in which medication was skipped and exposure to medication was insufficient for either IR or OROS.

Table 3 | Definition of model states

Health state	Definition	Medication intake per day:	
		OROS	IR
Optimal (A)[a]	Optimal[a] daily exposure to medication; remission[b] of ADHD symptoms; the child functions well with this treatment; no significant problems at home, at school, with peers or during leisure time; the child receives additional care such as visits to a specialist, behavioral therapy, extra attention at school, etc	1x	3x
Suboptimal (B)[c]	Insufficient daily exposure to medication; ADHD symptoms present but reduced; different from normal functioning; the child functions considerably well with this treatment; during short periods the child experiences problems at home, at school, with peers or during leisure time; the child receives additional care such as visits to a specialist, behavioral therapy, extra attention at school, etc	None	0-2x
Treatment stopped (C)	Treatment stopped in spite of remaining symptoms of ADHD; noticeable problems at home, at school, with peers and/or during leisure time; the child experiences more persistent hinder of those problems; the child receives additional care such as visits to a specialist, behavioral therapy, extra attention at school, etc	None	None
Remission (D)	No medication used; behavioral problems are no more different from normal; no more additional care needed related to ADHD such as visits to a specialist, behavioral therapy, extra attention at school, etc	None	None

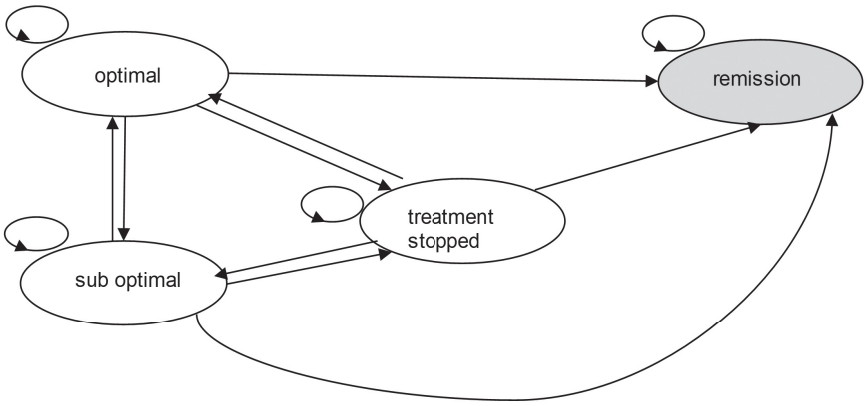
ADHD attention deficit hyperactivity disorder, IR immediate-release, OROS osmotic-release oral system
[a]Optimal intake is defined as follows: good compliance with intake of 1x/day for OROS and 3x/day for IR.
[b]Remission=not different from normal, symptoms of ADHD are at the most sometimes present, but not often or always.
[c]Suboptimal intake: insufficient compliance. Medication is not taken as prescribed, which means no intake for OROS and average intake of 1x/day for IR.

In the optimal state, patients were assumed to adhere to the prescribed medication and consequently not experience any symptoms of ADHD. Symptoms not directly related to ADHD but to comorbidity may still be present in this state. In a suboptimal state, in contrast, patients were assumed not to adhere properly to their prescribed medication, resulting in symptoms of ADHD and behavior different from normal behavior for their age group. As a single dose of OROS was required per day, skipping medication meant no medication at all in that state. For patients treated with IR, non-adherence at a mean of three prescribed doses per day [9] was assumed as either missing one, two or three doses per day yielding a mean of two missed doses per day in the suboptimal state.

Patients who stopped treatment entirely in spite of remaining symptoms of ADHD entered the state 'treatment stopped'. Patients with functional remission not needing medication for treatment of ADHD entered the state 'remission'. In line with the study performed by Faber et al. [29], we assumed that once in remission, patients remained in that state, which acted as an absorbing state (figure 1). The consulted psychiatrists indicated that reaching the state of remission would be exceptional. According to the experts the assumption of remission as an absorbing state could reasonably be made. However, the experts noted that there may be exceptions where patients experience a relapse after having reached the state of remission.

Patients in an 'optimal', 'suboptimal' or 'treatment stopped' state either remained in that state or transferred to one of the other states.

Figure 1 | Graphical representation of the model

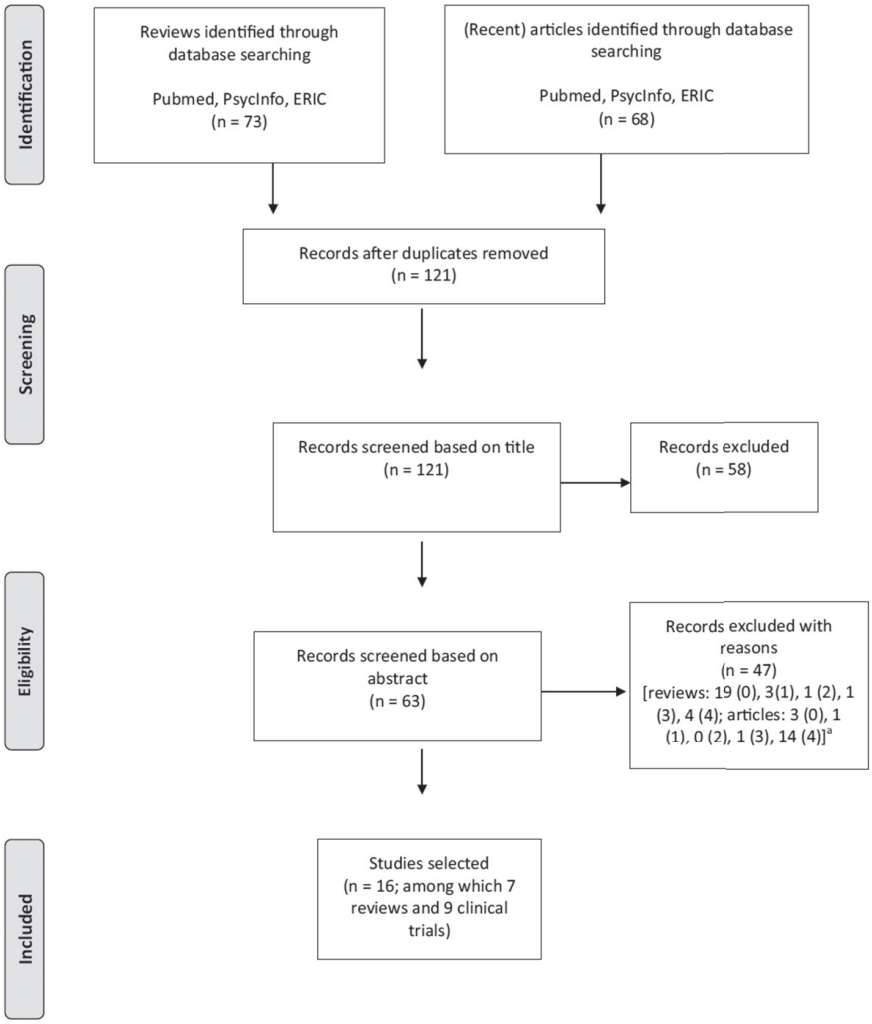


Transition rates

Variation in effect was modeled based on compliance and resulting symptom and behavior change (table 3). Data on transition rates between model states had to comply with our specific target population (i.e. children or adolescents with ADHD who initially had responded suboptimally to IR due to incorrect intake of medication). Furthermore, to guarantee the validity of model results, we preferred transition rates departing from one states to different states to all originate from the same source (and refer to the same definition of an optimal and suboptimal state). We considered transition rates from the Faber model [29] suboptimal as some of the transition rates were counterintuitive and the rates were based on multiple sources (i.e. literature and expert opinion). Hence, we performed a systematic literature review in the PubMed, PsycInfo and ERIC databases as to identify data to determine the transitions. First, we searched for reviews for the period from January 1, 2008 (the year of publication of the Faber model [29]) onwards. This search was performed on November 9, 2014. Then, we performed an additional search in the same databases directed at recent clinical trials from the publication date of the most recent identified review onwards. This second search was performed on December 8, 2104. Search terms for both searches were as follows: *ADHD OR "attention deficit hyperactivity disorder" [title] AND methylphenidate OR MPH OR MPH-IR OR MPH-ER OR pharmaco* [title] AND effect* OR efficacy OR cost-effectiveness OR cost-utility [title]*

The searches resulted in a total of 121 hits after duplicates were removed. The records were screened by two researchers independently, in a first round on title and in a second round on abstract. Where there was conflict, a decision was reached through consensus. The screening and selection process is summarized in a PRISMA flow diagram in figure 2.

Figure 2 | PRISMA flow diagram of systematic literature review



The selections based on title and abstract resulted in 16 studies to be included, among which were seven reviews and nine clinical trials. We were specifically interested in data from which transition rates for OROS and/or IR could be derived. Consultation of the reviews yielded several conclusions. Five reviews presented only mean scores on specific outcome measures [35, 36] or effect sizes [37-39]. Confidence intervals of effect sizes may be used to calculate transition rates based on a minimal meaningful improvement (i.e. defining a certain point on the distribution at which a patient moves from an optimal towards a suboptimal model state). However, as different underlying studies used

different outcomes as the basis for the stated effect sizes, minimal meaningful improvements (and hence, definition of the suboptimal model state) would differ per outcome measure and per study. Hence, we did not consider this approach a feasible option within the scope of this study. Another review presented information on pharmacokinetics [40]. One other study concerned a review of cost-effectiveness outcomes, not presenting specific data on state transitions [41]. None of the reviews specifically addressed the targeted patient population (i.e. children or adolescents who had initially responded suboptimally to IR due to incorrect intake of medication). Hence we considered the option to base transition estimates on a single study and examined the recent articles for further informative data.

From consultation of these articles we noted that seven of the nine articles did not contain suitable information. Two articles concerned letters to the editor [42, 43], one article was written in Iranian language [44], one articles concerned an explanatory study on effect sizes [45], one article presented mean scores [46] and one article referred to differences in scores [47]. Another article presented percentages of patients who improved (a potentially suitable measure for the calculation of transition probabilities). However, the study considered patients treated with specific extended-release MPH with 50% short-acting and 50% long-acting components [48]. Two remaining articles presented data potentially useful for calculation of transition rates [49, 50]. Garg et al. [49] found a treatment response of 90.7% in patients receiving IR (n=33) in Northern India. Soutullo et al. [50] stated that 51% (95% CI 31.1-60.6) of European patients (n=111) responded to treatment with OROS. The trial was performed in 48 centers across 10 European countries. However, both articles did not consider the specific patient population of this study and only one broad rate of response for the entire treatment period was provided, whereas our model included more specific transitions between the optimal and suboptimal states (back and forth) and accounted separately for patients staying in a specific state. Furthermore, Garg et al. [49] and Soutullo et al. [50] used different outcome measures to define response and the studies were performed in two different treatment populations. Hence, we considered the information available from these single clinical trials insufficient to use in the model. Consequently, we considered the consultation of an expert panel (from within the Dutch context) superior to using data from multiple international trials.

Hence, transition rate estimates were attained from consultation with a Delphi panel of experts (table 1). We retrieved all transition rate estimates from one consistent source (i.e. the expert panel).

The consulted psychiatrists suggested that the group of patients suboptimally treated with IR would, in particular, experience practical problems with accurate medication intake schemes during the day or at school. These patients would need to put more effort into adherence to the administration scheme compared with OROS, for which administration is limited to once a day. These differences in effect and effort were reflected in the transition rates between states.

Transition estimates were attained by blind questionnaires in two rounds, according to Dutch guidelines for pharmacoeconomic research [30] and consistent with the Delphi panel method [51, 52]. The experts were consulted independently and were not aware of the identity of the other experts joining the panel. Before distributing the questions to the experts, it was decided that consensus was supposed to be reached after two rounds of answers when (a) feedback of the experts was clear and (b) when experts did not all change their answers on the basis of the mean of the feedback of the first round. The questions for the panel were sent and returned by email. One of the researchers registered the replies anonymously. After all experts had returned the questionnaires, their answers were combined. The mean value for each question constituted the basis for the final answer to each question. The proposals for the final answers as well as the anonymized individual answers of the participants were reported to the experts after round 1. In the second round, experts were asked whether they intended to change their previous answers on the basis of the proposal for the final answer.

Utility values

ADHD is associated with reduced health-related quality of life [53-56]. The present model was built to assess the cost utility of OROS versus IR in children and adolescents with ADHD. Effects were expressed in terms of quality-adjusted life-years (QALYs). Several members of our research team were involved in a recent Dutch study that measured the quality of life of children with ADHD and their parents [57]. The study of van der Kolk et al. [57] was a cross-sectional study among member of a Dutch ADHD parent association. Data collection occurred via online questionnaires. The quality of life of the children (n=618) was based on parent proxy ratings, and the quality of life of the caregivers (n=590) was based

on self-report of the Euroqol (EQ-5D) questionnaire [57, 58]. The available quality of life data were highly suitable for inclusion in the current model as the state definitions of responders and non-responders closely matched the definition within the current model.

Utility of the patient

We found a significant difference in quality of life of patients compliant with prescribed medication compared with non-compliant patients [57]. Compliant patients reported a quality of life of 0.84 (ages 8-12 years 0.82; ages 13-18 years 0.86) whereas non-compliant patients reported a quality of life of 0.75 (ages 8-12 years 0.74; ages 13-18 years 0.77) [57]. In the current model, we included the quality of life values of the compliant group for the state 'optimal' and the utilities of the non-compliant group for the state of 'suboptimal' functioning. As there was no utility available for patients who had stopped treatment, we considered it reasonable to assign to those patients the same utility as patients in the suboptimal state, as this would constitute a conservative estimate. Based on the available data, utility was modeled to differ per model state but not per treatment type.

Spillover effects on caregiver

Family effects [8, 59-62] and negative effects of ADHD on families in particular [26, 63] have been addressed several times in the literature. Le et al. [8] suggest that benefits of ADHD treatment may also extend further than the individual patient. Brouwer et al. [19] proposed that when taking a societal perspective, these effects may be added to the effects experienced by patients. Hence, we considered it valuable to include spillover effects on the utility of a parent in the model. In our recent study on quality of life [57] we found a significant correlation between the quality of life of the child and the caregiver. No significant difference was found between the quality of life of parents of compliant or non-compliant children.

The literature on ADHD is very limited on this aspect, and our study [57] was the first study to report utilities of patients with ADHD and caregivers in one study. Further studies on the specific effect of ADHD on caregiver utility could not be retrieved from the literature. However, there is evidence available on the effect of a child with ADHD on health expenditures of caregivers. Hakkaart et al. [4] stated that 25% of health care expenditures of the caregiver of a child with ADHD can be attributed to the behavioral problems of the child. This suggests a

considerable influence of child health on caregiver health. In the absence of more specific data on the caregiver effects of ADHD, we searched for publications on caregiver effects in other diseases. Evidence from the literature on meningitis [23] suggests that 'spillover' health effects on family may constitute as much as 48% of the health effects on the patient. In the case of ADHD, this may be a conservative estimate as ADHD has been found to be especially stressful on the direct environment of the patient. Hence, as an estimate, we included 48% of caregiver utility in the model.

Cost parameters

Categories of direct medical and non-medical costs were kept consistent with the Faber model [29]. These categories were medication costs, costs of medical consultations, costs of medical and non-medical interventions, and costs of special education. Costs differed per state and in remission, we assumed no costs associated with ADHD. We assumed all costs except drug costs to be only dependent on the state and not on the type of medication (IR or OROS) received by the patient. This assumption was based on evidence from the literature on comparable efficacy of IR and OROS under the provision of full therapy compliance [10-12] and was confirmed by the expert panel of psychiatrists (table 1). We considered different costs for patients when below the age of 12 years and at and above the age of 12 years. This modeled difference in costs according to current age was based on consultation of the expert panel (table 1). The experts suggested differences in cost when switching schools (i.e. from primary to secondary education), which corresponds to the age of 12 years in the Dutch setting. Health care consumption (i.e., frequencies of consultations and non-pharmacological interventions) were extracted from the study performed by Faber et al. [29]. All costs were valued in Euros (2014). Cost prices were updated based on Hakkaart et al. [64], costs of special education were updated as reported by the Dutch Ministry of Education [65] and all costs were adjusted to 2014 values.

Next to the cost categories consistent with the Faber model [29], literature and available data of additional cost categories were searched to determine relevance and feasibility of inclusion in the model. Considered categories were: criminal justice costs, costs of lower-proficiency work and low income, out-of pocket expenses and spillover effects on caregivers (i.e., health care costs and production losses).

Costs of medication

Individuals in the OROS arm of the model used a full daily dose of OROS per day in the optimal state and no medication in all other states. In the IR arm of the study, participants were assumed to take the full daily dose of IR a day in the optimal state and on average 1/3 of the daily dose in the suboptimal state. The daily dose of both OROS and IR was determined on the basis of the average daily dose of two age groups (6-12 and 13-18 years) and was based on IMS data [66]. Cost of medication was based on the Dutch pricelist [67].

Costs of medical consultations

Consultation costs concerned contacts with psychiatrists, other medical specialists, general practitioners, and crisis contacts. The number of visits per year was dependent on age and based on the Faber model [29]. Unit prices were retrieved from the Dutch manual for costing research [64] and applied to the number of contacts.

Costs of medical and non-medical interventions

Intervention costs included costs of psychosocial and psychotherapeutic interventions as well as interventions for educational support (i.e., psycho education, parent training, behavior child therapy, social skills training, teacher training, remedial teaching, physical therapy, home training/care, outpatients' treatment and institutionalization). These categories were in line with the Dutch clinical guidelines for ADHD [9]. Interventions that are provided on a limited scale in the Netherlands (i.e., neurofeedback, cognitive training, mindfulness, diet) have not been included. The number of contacts was based on the Faber model [29]. Intervention costs were assumed to occur at age 6 and at age 12 for one year each as experts from the panel of consulted psychiatrists (table 1) indicated that those costs mainly occurred at the moment of switching between schools. Unit prices were retrieved from the Dutch manual for costing research [64] and applied to the number of contacts.

Costs of special education

Costs for special education were additional costs per day in special education. Advice for placement in special education was assumed dependent on age. Costs for special education were considered continuous from age 6 to age 18 in accordance with the experts' opinion. Probability of placement was based on the Faber model [29], and unit prices were based on the Ministry of Education, Culture and Science [65].

Criminal justice costs

Several authors have found a positive relationship between ADHD in childhood and antisocial behavior and drug use in (young) adults [68-70]. However, it has to be taken into account that the high degree of antisocial activity may be attributed to comorbid conduct disorder [71]. A recent study by Lichtenstein et al. [72] suggested that criminal behavior of ADHD patients decreases when medication is taken consistently. Evidence from the literature suggests that data on criminal justice costs related to ADHD are scarce [20] and especially limited in the European context [8]. Though these costs are considered highly relevant especially in the light of a possible relation with medication intake, the lack of available data resulted in the exclusion of these costs from the current model.

Costs for educational support, cost of lower-proficiency work and low income

Evidence from literature suggests that the impact of ADHD may exceed the age of school-going children and that it may result in poor educational performance [4, 8, 63, 73, 74], work achievements [75, 76] and household income [20, 70, 77, 78]. However, it is not yet clear whether medical treatment necessarily improves academic performance or income, as it may have an effect on some aspects of academic functioning and not on others [73]. Children with ADHD often require additional support within the educational setting [20]. As this study focused on children between 6 and 18 years, the costs of additional educational support within the education system up to age 18 were included within the cost categories 'costs of medical and non-medical interventions' and 'costs of special education' in the model (i.e., costs for teacher training, remedial teaching and costs of special education). When expanding current projections to a lifetime perspective, long-term consequences of educational effects (i.e., on work and income) should be included as well.

Out-of-pocket expenses

In a Dutch study on out-of-pocket expenses of children and adolescents with ADHD, Hakkaart et al. [4] presented data from parents of children with ADHD treated by a pediatrician. The authors found out-of-pocket expenses of 23.13 EUR (standard deviation EUR 150.35; adjusted to 2014 EUR) per annum in the Dutch setting. As the amount of out-of-pocket expenses is negligible (i.e., not significantly different from zero) in the study by Hakkaart et al. [4], we did not include these expenses in the current model.

Spill-over on caregivers (medical costs and production losses)
Hakkaart et al. [4] found that mean health care costs of mothers of children with ADHD were significantly higher than those of mothers of healthy children. Mean medical costs per year were 841.93 EUR (adjusted to 2014 EUR) for mothers of children with ADHD compared with 178.10 EUR of mothers of a healthy child. The authors stated that 25% of mothers noted that their use of health care services was related to the behavioral problems of their child [4]. Consequently, we assumed health care costs for a caregiver of 0,25 x (841.93 EUR-178.10 EUR) in the suboptimal and treatment stopped states and included these costs in the model. In the optimal state, no additional costs were assigned.

Hakkaart et al. [4] also collected data on production losses of mothers of patients with ADHD. The authors found significantly higher production losses in mothers of children with ADHD compared to mothers of healthy controls. Mean annual production losses of mothers (reduced efficiency and absence from work) were 2,594.03 EUR (adjusted to 2014 EUR) compared to 779.48 EUR for mothers of healthy children. As noted above, Hakkaart et al. [4] found that 25% of health care costs of the mother were related to behavioral problems of the child. It seems reasonable to assume that also 25% of production losses can be attributed to the behavioral problems of the child. Hence, in the model, we included mean annual production losses of 0.25 x (2,594.03 - 779.48 EUR) in the suboptimal and treatment stopped states. In the optimal state, no additional costs were assigned.

Model validation

Face validity was ascertained by consulting experts in the field of ADHD in the Netherlands on clinical aspects of model structure, model parameters and model input. Furthermore, verification of transition rates was attempted. Because of the scarce available data, we could only globally verify the number of patients in an optimal state after one year with response percentages from the literature identified from the systematic review [48-50], which we performed as part of the search for suitable transition rates. Though the estimates within these studies were based on different definitions of response or improvement and studies were performed in different countries, this constituted the best available data. As our study was performed in the population of patients who had in the past been treated with IR and reacted suboptimally because of problems with medication intake, it was expected that overall response within the existing literature would be higher than in our model. This rationale was supported, as Garg et al. [49] reported a 91% treatment response in patients treated with

MPH, Sobanski et al. [48] found 78% of patients receiving combined short- and long-acting MPH treatment had reduced symptoms and, according to Soutullo et al. [50], 51% of patients responded to treatment with OROS. On the basis of the expert panel estimates within the current model, 22% of patients treated with IR and 36% of patients treated with OROS achieved a transition from a suboptimal to an optimal state after one year. Hence, the transition estimates in our model appear to be in line with expectations and may even be conservative. We further performed scenario analyses to examine the sensitivity of model results to these parameters.

Sensitivity analyses

Sensitivity analyses were performed for four scenarios: one scenario assuming equal transition rates for IR and OROS; a second scenario including an augmented daily dose of exposure to medication; a third scenario excluding medical costs and production losses of the caregiver; and a forth scenario excluding the utility of caregivers. As transition rates were based on expert opinion (table 4), we performed a scenario to estimate the impact of these parameters on the results. Furthermore, due to issues of noncompliance, the daily dose data may provide an underestimation of optimal exposure. To measure the effect of this potential bias, a scenario was estimated which corrected for noncompliance. Studies by Adler and Nierenberg [16] and Swanson [79] have estimated noncompliance to amount to 13-64% and 20-65%, respectively. On the basis of these findings, the scenario considered an average of 40% noncompliance in daily dose data used (implying augmentation of the daily dose by 67% for both treatment arms). Two additional scenarios were performed to estimate the effect of the caregiver costs and effects on the model outcomes. As the underlying data for the inclusion of these model components was limited, the outcomes of the scenario analysis may provide further incentive for future data collections. One scenario was performed excluding medical costs and production losses of caregivers, and another scenario was performed where utilities of caregivers were excluded. Monte Carlo results were simulated per scenario, allowing for uncertainty around all parameter estimates while analyzing the specific effect of changes of the parameters of interest. Detailed model parameters are provided in table 4.

Table 4 | Detailed model parameters and assumptions [in EUR (2014)]

Parameter	Description	Source
General parameters		
Discount rate	Costs discounted at constant discount rate of 4%, effects at constant discount rate of 1,5%	College voor Zorgverzekeringen [30]
Patient age	All patients assumed to enter the model at age 6	Indicatie Concerta, Landelijke Stuurgroep Multidisciplinaire Richtlijnontwikkeling in de GGZ [9]
Monte Carlo random sampling	N=1000	Briggs et al. [33]
Transition probabilities		
IR: A to A	Dirichlet, mean 8.97	Expert panel data
IR: A to B	Dirichlet, mean 90.20	Expert panel data
IR: A to C	Dirichlet, mean 1.01	Expert panel data
IR: A to D	0	Expert panel data
IR: B to A	Dirichlet, mean 22.47	Expert panel data
IR: B to B	Dirichlet, mean 54.25	Expert panel data
IR: B to C	Dirichlet, mean 23.28	Expert panel data
IR: B to D	0	Expert panel data
IR: C to A	Dirichlet, mean 16.58	Expert panel data
IR: C to B	Dirichlet, mean 10.26	Expert panel data
IR: C to C	Dirichlet, mean 73.16	Expert panel data
IR: C to D	0	Expert panel data
OROS: A to A	Dirichlet, mean 6.25	Expert panel data
OROS: A to B	Dirichlet, mean 93.75	Expert panel data
OROS: A to C	0	Expert panel data
OROS: A to D	0	Expert panel data
OROS: B to A	Dirichlet, mean 58.91	Expert panel data
OROS: B to B	Dirichlet, mean 23.81	Expert panel data
OROS: B to C	Dirichlet, mean 17.27	Expert panel data
OROS: B to D	0	Expert panel data
OROS: C to A	Dirichlet, mean 24.21	Expert panel data
OROS: C to B	Dirichlet, mean 14.21	Expert panel data
OROS: C to C	Dirichlet, mean 61.58	Expert panel data
OROS: C to D	0	Expert panel data
Utility – patient (8-12 years)		
Optimal	Beta, mean 0.82, se 0.0979	van der Kolk et al. [57]
Suboptimal	Beta, mean 0.74, se 0.01588	van der Kolk et al. [57]
Treatment stopped	Beta, mean 0.74, se 0.01588	van der Kolk et al. [57]
Utility – patient (13-18 years)		
Optimal	Beta, mean 0.86, se 0.01097	van der Kolk et al. [57]
Suboptimal	Beta, mean 0.77, se 0.02645	van der Kolk et al. [57]
Treatment stopped	Beta, mean 0.77, se 0.02645	van der Kolk et al. [57]

Table 4 | Continued

Parameter	Description			Source
Utility - caregiver				
Optimal	Beta, mean 0.85, se 0.00897			van der Kolk et al. [57]
Suboptimal	Beta, mean 0.83, se 0.01499			van der Kolk et al. [57]
Treatment stopped	Beta, mean 0.83, se 0.01499			van der Kolk et al. [57]
Drug costs				
Daily dose OROS – child 6-12 years (mg)	31.70			IMS Health BV [66]
Daily dose OROS – child 13-18 years (mg)	39.10			IMS Health BV [66]
Daily dose IR – child 6-12 years (mg)	22.00			IMS Health BV [66]
Daily dose IR – child 13-18 years (mg)	29.20			IMS Health BV [66]
Costs/ mg OROS	0.05			Zorginstituut Nederland [67]
Costs/ mg IR	0.01			Zorginstituut Nederland [67]
Pharmacy fee/ 3 months	7.0			Zorginstituut Nederland [67]
Consultation costs	Incurred by children between 6 and 18 years			
Number of visits per year – child < =12	State A	State B	State C	
Psychiatrist	2.28	3.42	5.00	Faber et al. [29]
Other specialist	0	0	1.38	Faber et al. [29]
General Practitioner	0	0	0.58	Faber et al. [29]
Crisis contacts	0.57	1.49	2.71	Faber et al. [29]
Number of visits per year – child > 12	State A	State B	State C	
Psychiatrist	2.43	3.57	5.00	Faber et al. [29]
Other specialist	0	0	0.11	Faber et al. [29]
General Practitioner	0	0.29	0.43	Faber et al. [29]
Crisis contacts	0.35	1.28	3.00	Faber et al. [29]
Costs per visit				
Psychiatrist	113.53			Hakkaart et al. [64]
Other specialist	75.15			Weighted average psychiatrist and medical specialist: 46:34 [29, 64]
General Practitioner	31.22			Hakkaart et al. [64]
Crisis contacts	256.20			Based on Tariffs AWBZ-institutions 2005 [88]
Intervention costs	Incurred by children of age 6 and of age 12.			
Transferred % of patients – child < =12	State A	State B	State C	
Psycho education	0.89	0.93	1.00	Faber et al. [29]
Parent training	0.49	0.76	0.79	Faber et al. [29]
Behavior therapy child	0.07	0.23	0.57	Faber et al. [29]

Table 4 | Continued

Parameter	Description			Source
Social skills training (SOVA)	0.19	0.28	0.38	Faber et al. [29]
Teacher training	0.43	0.57	0.66	Faber et al. [29]
Remedial teaching	0.37	0.51	0.77	Faber et al. [29]
Physical therapy	0	0	0	Faber et al. [29]
Home training/care	0.04	0.13	0.33	Faber et al. [29]
Outpatients' treatment	0	0	0.25	Faber et al. [29]
Institutionalization	0	0	0.03	Faber et al. [29]
Transferred % of patients – child >12	State A	State B	State C	
Psycho education	0.94	0.90	0.89	Faber et al. [29]
Parent training	0.31	0.44	0.74	Faber et al. [29]
Behavior therapy child	0.09	0.28	0.56	Faber et al. [29]
Social skills training (SOVA)	0.07	0.26	0.53	Faber et al. [29]
Teacher training	0.10	0.33	0.32	Faber et al. [29]
Remedial teaching	0.02	0.39	0.47	Faber et al. [29]
Physical therapy	0	0	0	Faber et al. [29]
Home training/care	0	0.1	0.13	Faber et al. [29]
Outpatients' treatment	0	0	0.26	Faber et al. [29]
Institutionalization	0	0	0.04	Faber et al. [29]
Number of visits per year – child < =12	State A	State B	State C	
Psycho education	2.64	3.64	3.86	Faber et al. [29]
Parent training	8.34	7.92	14.01	Faber et al. [29]
Behavior therapy child	13.18	11.80	13.15	Faber et al. [29]
Social skills training (SOVA)	9.15	9.79	9.15	Faber et al. [29]
Teacher training	1.75	3.73	3.94	Faber et al. [29]
Remedial teaching	20.00	20.00	20.00	Faber et al. [29]
Physical therapy	6.00	0	0	Faber et al. [29]
Home training/care	10.00	11.15	14.31	Faber et al. [29]
Outpatients' treatment	0	0	51.75	Faber et al. [29]
Institutionalization	0	0	90.00	Faber et al. [29]
Number of visits per year – child > 12	State A	State B	State C	
Psycho education	2.78	3.57	5.42	Faber et al. [29]
Parent training	5.91	8.24	13.74	Faber et al. [29]
Behavior therapy child	10.00	11.44	12.88	Faber et al. [29]
Social skills training (SOVA)	9.15	11.44	10.59	Faber et al. [29]
Teacher training	2.00	2.50	3.73	Faber et al. [29]
Remedial teaching	20.00	20.00	20.00	Faber et al. [29]
Physical therapy	0	0	0	Faber et al. [29]

Table 4 | Continued

Parameter	Description			Source
Home training/care	0	10.00	10.06	Faber et al. [29]
Outpatients' treatment	0	0	51.75	Faber et al. [29]
Institutionalization	0	0	135.00	Faber et al. [29]
Costs per visit				
Psycho education	111.17			Based on Tariffs AWBZ-institutions 2005 [88]
Parent training	104.15			Based on Tariffs AWBZ-institutions 2005 [88]
Behavior therapy child	111.17			Based on Tariffs AWBZ-institutions 2005 [88]
Social skills training (SOVA)	111.17			Based on Tariffs AWBZ-institutions 2005 [88]
Teacher training	76.05			Based on Tariffs AWBZ-institutions 2005 [88]
Remedial teaching	58.49			Based on Dutch Society of Remedial Teachers [89]
Physical therapy	39.84			Hakkaart et al. [64]
Home training/care	114.52			Based on Health care insurance board [90]
Outpatients' treatment	150.57			Hakkaart et al. [64]
Institutionalization	301.09			Hakkaart et al. [64]
Special education costs	Incurred by children between 6 and 18 years			
	State A	State B	State C	
Advice placement special education (%) – child< =12	0.015	0.1224	0.4356	Faber et al. [29]
Advice placement special education (%) – child> 12	0.0007	0.0863	0.3711	Faber et al. [29]
Additional costs special education/day	13.63			Based on Ministry of Education, Culture and Science [65]

IR immediate-release, OROS osmotic-release oral system, SE standard error

Results

Transition estimates

In accordance with the model population, the expert panel of psychiatrists (table 1) estimated transition rates for a patient population initially treated with IR with suboptimal result because of incorrect intake of medication.

Table 5 displays mean transition percentages per day as estimated by the expert panel.

Table 5 | Mean (standard deviation) transitions per day (in %) as estimated by expert panel

From/to	optimal	suboptimal	treatment stopped	remission
IR				
optimal	8.79 (6.34)	90.20 (7.84)	1.01 (2.02)	0 (0)
suboptimal	22.47 (21.41)	54.25 (14.21)	23.28 (11.74)	0 (0)
treatment stopped	16.58 (8.02)	10.26 (8.60)	73.16 (15.94)	0 (0)
OROS				
optimal	6.25 (9.46)	93.75 (9.46)	0 (0)	0 (0)
suboptimal	58.91 (21.03)	23.81 (12.51)	17.27 (9.82)	0 (0)
treatment stopped	24.21 (11.73)	14.21 (15.81)	61.58 (23.98)	0 (0)

IR immediate-release, *OROS* osmotic-release oral system

Variability in cost parameters was captured by gamma distributions around the mean, and variability in transition probabilities entered the model through Dirichlet distributions [33]. Samples from these prior distributions were drawn by Monte Carlo simulation. For illustrative purposes and in the absence of trial data, as a common simplifying assumption, the standard errors of the cost parameters were assumed 20% of the mean. As is common in probabilistic models, a total of 1000 Monte Carlo simulations were performed to generate the model results.

Different estimates were attained for patients receiving OROS and for patients receiving IR. Experts estimated the probability to transfer from a suboptimal or treatment stopped state to an optimal state to be higher for patients receiving OROS than for patients receiving IR. However, they predicted patients receiving OROS to have a lower chance of staying in an optimal state than patients receiving IR. Furthermore, the experts estimated patients receiving OROS to have a lower chance than patients receiving IR to stop treatment and a higher

chance to go back to an optimal or suboptimal state when having stopped the treatment. Transitions from the suboptimal state to the remaining states appear to differ most between treatments. All experts considered the transition to a state of remission to be 0% per day. This means that the state 'remission' becomes redundant. In line with earlier critical comments on possible relapse after remission, it becomes clear that remission is seen as an exceptionally rare state (Figure 3) such that patients are expected to keep moving between the optimal, suboptimal and treatment stopped states instead of reaching a stable state of remission.

Figure 3 | Expert comments on state of remission

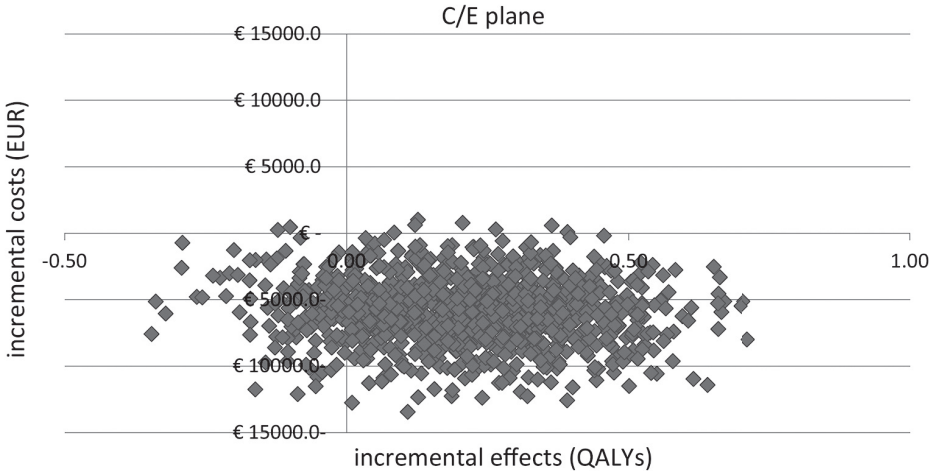
"I usually aim at stopping medication after approximately 2 years of treatment and starting from that moment once a patient has been well controlled for approximately 3 years, one can speak of remission."
"Even in a state of remission, children who suffer from ADHD will still experience long-term effects of ADHD such as impact on emotional development." "In general there are few children who actually achieve remission as I work with children between 0 and 18 years of age…One could question whether it is actually possible to achieve complete remission of ADHD for a child."
"Remission is a state in which a patient has no complaints over a certain time horizon. Often this time horizon is considered to be half a year."
"The aim of treatment is not to let symptoms disappear completely but to keep them under control. Total disappearance of symptoms is actually not possible…that would mean that a child is cured and that is very rare and is achievable over a timeframe of several years."

Model results and sensitivity analyses

Model results indicate dominance of OROS compared with IR in this population. OROS results in incremental QALY gains while saving costs. The number of QALYs for OROS exceeds the number of QALYs for IR by 0.22 (95% CI -0.206, 0.228), and the total costs of OROS are estimated to be lower than IR with incremental cost savings of 5,815 EUR (95% CI 5,661 EUR, 5,969 EUR) (table 6). These results suggest that, for this patient group, OROS produces better effects at lower cost compared with IR. The detailed probabilistic model results of 1000 Monte Carlo simulations are presented on a cost-effectiveness (C/E) plane and as a cost-effectiveness acceptability curve (CEAC) [80] in figures 4 and 5, respectively. Figure 4 provides details on the uncertainty around the costs and effect of OROS compared with IR. The 1000 points in the scatter plot each

represent one simulation result. The *x*-axis displays the amount of incremental QALY gains or losses and the *y*-axis shows the incremental costs expressed in Euros (EUR).

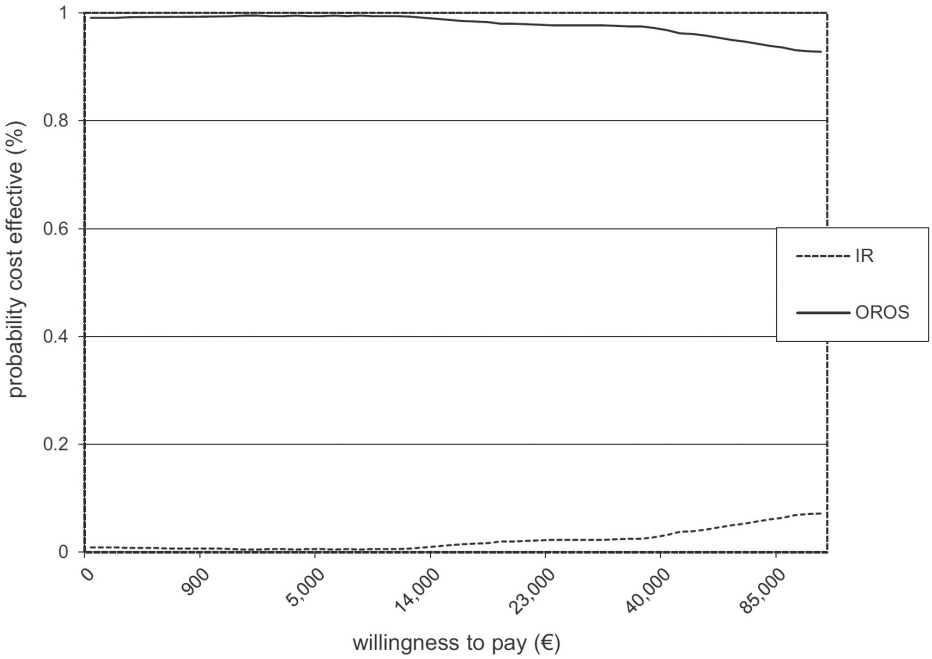
Figure 4 | Scatter plot: incremental costs and effects based on 1000 Monte Carlo simulations



The results of the C/E plane show that the majority of data points appear in the southeast quadrant, with lower costs and higher effects of OROS compared with IR, which indicates dominance of OROS versus IR.

Figure 5 shows a graphical presentation of the CEAC, displaying the probability that OROS is cost-effective compared to IR given different values of maximum threshold for society. The threshold values in terms of Euros are shown on the x-axis and the probability of OROS being cost-effective is displayed on the y-axis.

Figure 5 | CEAC: probability of OROS being cost-effective compared with IR



The CEAC displays data points within all four quadrants, with the majority of data points in the southeast quadrant [81]. The probability of OROS being cost-effective ranges between 93 and 99%. The CEAC does not cross the *y*-axis at 0 as some data points in the C/E plane display cost savings of OROS compared with IR. Furthermore, the CEAC does not asymptote to 1 because a part of the observed data points on the C/E plane show negative incremental effects.

Sensitivity analyses indicate that when transition rates of OROS are equal to the transitions of IR (scenario 1), the incremental QALYs gained for OROS compared to IR amount to 0.00 and costs of treatment with OROS appear slightly higher than treatment with IR, with additional incremental costs of 800 EUR (table 6). These results in terms of incremental costs are close to the results of the Faber model [29] and the incremental effects are reduced to zero. Model results, thus, are strongly dependent on accurate estimates of transition probabilities as these determine how fast patients move between model states and, hence how often they stay in more or less 'expensive' states.

The scenario with an augmented daily dose of exposure to medication (scenario 2) resulted in slightly lower savings compared with the base scenario (savings

of 4,502 EUR and QALY gains of 0.21).Exclusion of medical costs and production losses of caregivers from the model results in incremental cost savings of 4,930 EUR and incremental QALY gains of 0.22.

When utility of caregivers is excluded, there are insignificant changes in incremental costs (5,900 EUR) and a decrease in incremental QALY gains to 0.15. This incremental QALY decrease is entirely explained by the fact that 48% of the QALYs of the patient have been added to account for effect on a caregiver.

Table 6 | Mean model results and sensitivity analyses of Monte Carlo simulations (*n*=1000)

	Description	Incremental costs	Incremental QALYs
Base case		-5,815	0.22
Scenario 1	Transition rates of OROS equal to transition rates of IR	800	0.00
Scenario 2	Daily dose of medication +66%	-4,502	0.21
Scenario 3	Medical costs and production losses caregiver excluded	-4,930	0.22
Scenario 4	Utility of caregivers excluded	-5,900	0.15

Discussion and Conclusions

Policymakers increasingly use cost-effectiveness analyses to inform decision-making on competing health care interventions. Health economic models facilitate these analyses by providing a framework to combine information from different sources and enable probabilistic estimations. Within health care there has been debate on which perspective to be taken in such models. In the Netherlands, a societal perspective is common according to the health economic guidelines. Lately, there have been voices to even include broader effects (i.e., exceeding the patient and exceeding health care) [19]. This may be especially relevant to illness in children and disorders, which have a high impact on third parties (such as in the setting of ADHD). The presented model adds to the current movement towards broader considerations in cost-effectiveness analyses. We have presented a model compliant with the current health economic guidelines and at the same time considered, and where possible included, broader societal aspects to increase the comprehensiveness of the model results. Hence, the results of this study can be used as direct input to policymakers' decision making.

Model results indicate that, for children responding suboptimally to treatment with IR, the beneficial effect of OROS on compliance may be worth the additional medication costs. The current model was based on the Faber model [29] but model structure and input were improved and the model was enhanced with additional broader societal parameters. Transition rates consistent with our model structure could not be obtained from one source of literature. Therefore, we chose to consult an expert panel to provide transition rate estimates for all model states. Guidelines for health economic analysis state that in case where data are not available, the use of input from an expert panel is accepted, provided that a scientific method is used. The experts were consulted using a Delphi method as described in the Dutch guidelines for pharmacoeconomic analyses [30]. In case of transition probabilities, the use of an expert panel was crucial and far from ideal. However, it was necessary since data were not available from literature. We attempted to validate the expert transitions; however, because of the scarce literature, this was only partly possible. Hence, a scenario analysis was performed to examine sensitivity of model results to these parameters. This analysis showed that model results are very sensitive to estimated transitions. Hence, empirical data to improve these estimates are strongly needed. As elaborated in the "Methods" section, we adhered to the formal requirements of the Delphi method and present our results with caution as the focus of this study was to build an up-to-date model for evaluation of OROS compared with

IR rather than to gather comprehensive input to the model. One should note that participating experts received a small compensation, which was strictly limited to compensation for their invested time. The authors consider the collection of empirical data the necessary next step for further research.

In the current model, important societal costs and effect have been included (i.e., health care costs of caregiver, production losses of the caregiver and utilities of the caregiver). However, several aspects could not be covered, because of limited availability of data. Estimates of justice costs (i.e., incarceration costs, victim costs, etc.) could not be included, because of a lack of data in the considered age group and in the European context. Out-of-pocket expenses were not included as there was evidence from the Dutch literature that these costs were negligible [4]. However, when applying the current model to a different setting, country adaptation may be necessary, as a Belgian study [63] suggests differences in amounts of out-of-pocket expenses. These differences may be attributed to differences in sampling methods between the studies, but differences in health care systems may also play a role and necessitate model adaptation. Long-term effects on work and income have not been included in the model but are considered relevant. When taking a long-term perspective, these costs should be included in the model. Furthermore, in the current model, health care costs and production losses of the caregiver, which should be attributed to the behavioral problems of the child, were estimated at 25% of the total production losses. Mothers of children with ADHD had indicated that this percentage of their health care expenditures was related to the behavioral problems of their child. As it might be ethically difficult for a mother to blame her child for her medical problems, this estimate may be conservative. On the other hand, heritability of ADHD may point towards high medical expenditures of mothers for their own medical needs. Hence, in total the chosen percentage might be a good estimate. Concerning the utilities of caregivers, the literature was especially limited. Hence, additional data are necessary to provide a better basis for future analyses. So, especially, concerning societal costs, available data were extremely scarce, and we emphasize the necessity for additional studies to close this gap.

Earlier cost-effectiveness results of the Faber model [29] resulted in incremental costs of 276 EUR and incremental QALY gains of 0.13 of OROS compared to IR. The calculations in the current study resulted in mean incremental cost savings of 5,018 EUR and mean incremental QALY gains of 0.22. The differences in costs can be explained by substantial revision of transition rates based on expert panel

estimates, differences in model structure (e.g., consistent model states over treatment alternatives and omission of the state of remission because of experts' opinion), update of cost parameters to 2014 values (including a slight price deflation on drug costs of OROS) and difference in time horizon. Compared to the transition rates presented by Faber et al. [29], which were based on a collection of different sources, the experts' transition rates based on the expert panel presented in this study showed significant differences. In the Faber model [29] no differentiation was made between the probability of patients in an optimal or in a suboptimal state to stop treatment (IR treatment arm). The same was true for the probability of patients in the optimal or suboptimal state to achieve functional remission or to transfer to a non-compliant/suboptimal state (OROS treatment arm). With respect to these probabilities, the Faber model [29] treated patients in an optimal and suboptimal state as being equal. These assumptions appear rather strong as they a priori prevented differences in compliance affecting the chance of functional remission or termination of treatment. In this study, on the contrary, the experts indicated clear differences between these transition probabilities (table 5). Faber et al. [29] furthermore assumed a chance of moving from an optimal to a suboptimal state in the IR treatment arm to be 0 and the chance to stop treatment when non-compliant as 0 in the OROS arm. Both these assumptions appear counter-intuitive as they imply that a patient treated with IR may not miss a dose once he has achieved an optimal state and a non-compliant patient receiving OROS may not stop treatment at all. In the current study, the expert panel estimated all transition rates without assumptions beforehand to achieve a consistent framework of transition probabilities. As the transitions have a direct effect on how long patients remain in a state, these estimates have a strong influence on both incremental costs and effects and mainly explain the difference between the current model outcomes and those of the Faber model [29]. The transition estimates by the expert panel showed that patients receiving OROS were expected to be less likely to stop treatment, which corresponds to the findings from the literature on treatment duration and continuity [14, 15]. Furthermore, more patients receiving OROS were expected to move from a suboptimal state or a state were treatment was stopped back to an optimal state compared to patients receiving IR. These results are in line with the literature, as this suggests improved compliance of patients receiving OROS compared with those receiving IR [16-18]. However, the expert panel predicted the patients treated with OROS to have a lower probability to remain in an optimal state than patients receiving IR. This constitutes an unexpected finding given the

literature on better compliance of patients receiving OROS [16-18]; hence, we consider collection of additional empirical data necessary.

The treatment effects in terms of quality of life were based on parental preferences. These were taken from an existing study by van der Kolk et al. [57]. It has been shown that the value of quality of life valuation by children themselves may be questionable, particularly because of lack of language, cognitive limitations, long-term perspective [82, 83] and conceptual difficulty of the standard gamble task [84]. Therefore, parental preferences were considered most appropriate for the young population of the current study. However, we are aware of the shortcomings of this approach, namely the inability of parents to accurately estimate invisible and subjective aspects of their child's quality of life, such as social and emotional functioning [82-85]. This limitation may lead to inaccurate estimates and a possible overestimation of the child's disability [86]. Besides the utility of the patient, we also included caregiver utility in the model. The proportion included was based on very limited evidence from the literature. As a scenario analysis showed, model results are sensitive to these utilities. Hence, additional data are needed. As economic cost-utility analysis and expressing outcomes in terms of QALYs is not yet common in the field of child adolescent mental health [87], it may remain relevant for further studies to investigate results based both on costs/QALY and on costs per different (more clinically focused) outcome measures.

Severity of ADHD was not specified, but average severity was assumed in the model. One could, for instance, specify severity of ADHD in the model by distinguishing between the following categories of severity: 1) severe: no remission achievable; 2) moderate: 50% remission achievable; and 3) mild: 100% remission achievable. Furthermore, taking account of co-morbidities may affect the costs entered in the model in such a way that part of the costs, e.g., special education, may not be attributed to ADHD alone but to behaviors which arise from a combination of co-morbidities. In addition, long-term learning delay and emotional development problems were not taken into account in the current model. Hence, it should be noted that the current methodology is yet incomplete and that consideration of additions for long-term effects or consideration of different types of outcome measures to better account for the specific effects of mental health interventions may be necessary to improve the existing methodology.

Finally, as the focus of the study was the construction and demonstration of a broad and up-to-date probabilistic model compliant with current health economic methodology in this population rather than the provision of extensive input to the model, the model results should be interpreted with care. As can be seen from our results, future research should especially be directed at the collection of empirical data on transition estimates. We specifically suggest the collection of data from observational studies with large numbers of ADHD patients receiving (different types of) pharmaceutical treatment(s) compared with a control group of ADHD patients not treated with medication. It would be especially valuable to obtain data on treatment response (i.e., transition rates), health care use, school absence and performance, criminal activities and quality of life for these groups. To better cover the broad societal aspects, a very valuable and relevant addition would be data on the medical consumption, absence from work and utility of the caregivers and siblings as well. This information would be a valuable and necessary addition to the current model as it would lead to an increase in accuracy of the results and form a valuable basis for clinical and policy recommendations.

References

[1] Visser SN, Danielson ML, Bitsko RH, Holbrook JR, Kogan MD, Ghandour RM, et al. Trends in the parent-report of health care provider-diagnosed and medicated attention-deficit/hyperactivity disorder: United States, 2003-2011. J Am Acad Child Adolesc Psychiatry. 2014;53(1):34-46.e2.

[2] van den Ban E, Souverein P, Swaab H, van Engeland H, Heerdink R, Egberts T. Trends in incidence and characteristics of children, adolescents, and adults initiating immediate- or extended-release methylphenidate or atomoxetine in the Netherlands during 2001-2006. J Child Adolesc Psychopharmacol. 2010;20(1):55-61.

[3] Wicks-Nelson R, Israel AC. Abnormal Child and Adolescent Psychology. 8th ed. New York: Pearson; 2013.

[4] Hakkaart-van Roijen L, Zwirs BW, Bouwmans C, Tan SS, Schulpen TW, Vlasveld L, et al. Societal costs and quality of life of children suffering from attention deficient hyperactivity disorder (ADHD). Eur Child Adolesc Psychiatry. 2007;16(5):316-326.

[5] Lara C, Fayyad J, de Graaf R, Kessler RC, Aguilar-Gaxiola S, Angermeyer M, et al. Childhood predictors of adult attention-deficit/hyperactivity disorder: results from the World Health Organization World Mental Health Survey Initiative. Biol Psychiatry. 2009;65(1):46-54.

[6] Tuithof M, Ten Have M, van Dorsselaer S, De Graaf R. Prevalentie, persistentie en gevolgen van ADHD in de Nederlandse volwassen bevolking. Tijdschrift voor Psychiatrie. 2014;56:10--19.

[7] Doshi JA, Hodgkins P, Kahle J, Sikirica V, Cangelosi MJ, Setyawan J, et al. Economic impact of childhood and adult attention-deficit/hyperactivity disorder in the United States. J Am Acad Child Adolesc Psychiatry. 2012;51(10):990-1002.e2.

[8] Le HH, Hodgkins P, Postma MJ, Kahle J, Sikirica V, Setyawan J, et al. Economic impact of childhood/adolescent ADHD in a European setting: the Netherlands as a reference case. Eur Child Adolesc Psychiatry. 2013.

[9] Landelijke Stuurgroep Multidisciplinaire Richtlijnontwikkeling in de GGZ. Multidisciplinaire richtlijn ADHD - Richtlijn voor de diagnostiek en behandeling van ADHD bij kinderen en jeugdigen. 2005.

[10] Wolraich ML, Greenhill LL, Pelham W, Swanson J, Wilens T, Palumbo D, et al. Randomized, controlled trial of oros methylphenidate once a day in children with attention-deficit/hyperactivity disorder. Pediatrics. 2001;108(4):883-892.

[11] Pelham WE, Gnagy EM, Burrows-Maclean L, Williams A, Fabiano GA, Morrisey SM, et al. Once-a-day Concerta methylphenidate versus three-times-daily methylphenidate in laboratory and natural settings. Pediatrics. 2001;107(6):E105.

[12] Biederman J, Mick EO, Surman C, Doyle R, Hammerness P, Michel E, et al. Comparative acute efficacy and tolerability of OROS and immediate release formulations of methylphenidate in the treatment of adults with attention-deficit/hyperactivity disorder. BMC Psychiatry. 2007;7:49.

[13] Claxton AJ, Cramer J, Pierce C. A systematic review of the associations between dose regimens and medication compliance. Clin Ther. 2001;23(8):1296-1310.

[14] Kemner JE, Lage MJ. Impact of methylphenidate formulation on treatment patterns and hospitalizations: a retrospective analysis. Ann Gen Psychiatry. 2006;5:5.

[15] Marcus SC, Wan GJ, Kemner JE, Olfson M. Continuity of methylphenidate treatment for attention-deficit/hyperactivity disorder. Arch Pediatr Adolesc Med. 2005;159(6):572-578.

[16] Adler LD, Nierenberg AA. Review of medication adherence in children and adults with ADHD. Postgrad Med. 2010;122(1):184-191.

[17] Christensen L, Sasane R, Hodgkins P, Harley C, Tetali S. Pharmacological treatment patterns among patients with attention-deficit/hyperactivity disorder: retrospective claims-based analysis of a managed care population. Curr Med Res Opin. 2010;26(4):977-989.

[18] Gau SS, Shen HY, Chou MC, Tang CS, Chiu YN, Gau CS. Determinants of adherence to methylphenidate and the impact of poor adherence on maternal and family measures. J Child Adolesc Psychopharmacol. 2006;16(3):286-297.

[19] Brouwer W, van Exel J, Tilford M. Incorporating caregiver and family effects in economic evaluations of child health. In: Ungar W, editor. Economic Evaluation in Child Health: Oxford University Press; 2010.

[20] Beecham J. Annual research review: Child and adolescent mental health interventions: a review of progress in economic studies across different disorders. J Child Psychol Psychiatry. 2014;55(6):714-732.

[21] Bernfort L, Nordfeldt S, Persson J. ADHD from a socio-economic perspective. Acta Paediatr. 2008;97(2):239-245.

[22] Wu EQ, Hodgkins P, Ben-Hamadi R, Setyawan J, Xie J, Sikirica V, et al. Cost effectiveness of pharmacotherapies for attention-deficit hyperactivity disorder: a systematic literature review. CNS Drugs. 2012;26(7):581-600.

[23] Al-Janabi H, Van Exel J, Brouwer W, Trotter C, Glennie L, Hannigan L, et al. The family impact of meningitis: a cross-sectional analysis of long-term outcomes for survivors' family networks. to be submitted.

[24] King. S, Griffin S, Hodges Z, Weatherly H, Asseburg C, Richardson G, et al. A systematic review and economic model of the effectiveness and cost-effectiveness of methylphenidate, dexamfetamine and atomoxetine for the treatment of attention deficit hyperactivity disorder in children and adolescents. Health Technology Assessment. 2006;10(23).

[25] Marchetti A, Magar R, Lau H, Murphy EL, Jensen PS, Conners CK, et al. Pharmacotherapies for attention-deficit/hyperactivity disorder: expected-cost analysis. Clin Ther. 2001;23(11):1904-1921.

[26] Pelham WE, Foster EM, Robb JA. The economic impact of attention-deficit/hyperactivity disorder in children and adolescents. J Pediatr Psychol. 2007;32(6):711-727.

[27] Cottrell S, Tilden D, Robinson P, Bae J, Arellano J, Edgell E, et al. A modeled economic evaluation comparing atomoxetine with stimulant therapy in the treatment of children with attention-deficit/hyperactivity disorder in the United Kingdom. Value Health. 2008;11(3):376-388.

[28] Hong J, Dilla T, Arellano J. A modelled economic evaluation comparing atomoxetine with methylphenidate in the treatment of children with attention-deficit/hyperactivity disorder in Spain. BMC Psychiatry. 2009;9:15-244X-9-15.

[29] Faber A, van Agthoven M, Kalverdijk LJ, Tobi H, de Jong-van den Berg LT, Annemans L, et al. Long-acting methylphenidate-OROS in youths with attention-deficit hyperactivity disorder suboptimally controlled with immediate-release methylphenidate: a study of cost effectiveness in The Netherlands. CNS Drugs. 2008;22(2):157-170.

[30] College voor Zorgverzekeringen. *Dutch guidelines for pharmacoeconomic research, revised edition*. Diemen: College voor Zorgverzekeringen; 2006.

[31] Jonsson B. Ten arguments for a societal perspective in the economic evaluation of medical innovations. Eur J Health Econ. 2009;10:357--359.

[32] Secnik K, Matza LS, Cottrell S, Edgell E, Tilden D, Mannix S. Health state utilities for childhood attention-deficit/hyperactivity disorder based on parent preferences in the United Kingdom. Med Decis Making. 2005;25(1):56-70.

[33] Briggs A, Claxton K, Sculpher M. Decision Modelling for Health Economic Evaluation. New York: Oxford University Press; 2006.

[34] Jensen PS, Hinshaw SP, Swanson JM, Greenhill LL, Conners CK, Arnold LE, et al. Findings from the NIMH Multimodal Treatment Study of ADHD (MTA): implications and applications for primary care providers. J Dev Behav Pediatr. 2001;22(1):60-73.

[35] Coghill D, Banaschewski T, Zuddas A, Pelaz A, Gagliano A, Doepfner M. Long-acting methylphenidate formulations in the treatment of attention-deficit/hyperactivity disorder: a systematic review of head-to-head studies. BMC Psychiatry. 2013;13:237-244X-13-237.

[36] Armstrong RB, Damaraju CV, Ascher S, Schwarzman L, O'Neill J, Starr HL. Time course of treatment effect of OROS(R) methylphenidate in children with ADHD. J Atten Disord. 2012;16(8):697-705.

[37] Pelz R, Banaschewski T, Becker K. Methylphenidate of retard forms in children and adolescents with ADHD - an overview. Klin Padiatr. 2008;220(2):93-100.

[38] Wigal SB. Efficacy and safety limitations of attention-deficit hyperactivity disorder pharmacotherapy in children and adults. CNS Drugs. 2009;23 Suppl 1:21-31.

[39] Hodgkins P, Shaw M, Coghill D, Hechtman L. Amfetamine and methylphenidate medications for attention-deficit/hyperactivity disorder: complementary treatment options. Eur Child Adolesc Psychiatry. 2012;21(9):477-492.

[40] Maldonado R. Comparison of the pharmacokinetics and clinical efficacy of new extended-release formulations of methylphenidate. Expert Opin Drug Metab Toxicol. 2013;9(8):1001-1014.

[41] Catala-Lopez F, Ridao M, Sanfelix-Gimeno G, Peiro S. Cost-effectiveness of pharmacological treatment of attention deficit hyperactivity disorder in children and adolescents: qualitative synthesis of scientific evidence. Rev Psiquiatr Salud Ment. 2013;6(4):168-177.

[42] Jerome L. Effects of methylphenidate on acute Math performance in children with ADHD. Can J Psychiatry. 2014;59(5):290.

[43] Qi R, Zhang LJ, Lu GM. Emphasize the effect of methylphenidate on brain function in attention-deficit/hyperactivity disorder research. JAMA Psychiatry. 2014;71(2):210.

[44] Shahrbabaki ME, Sabzevari L, Haghdoost A, Ashtiani R. A randomized double blind crossover study on the effectiveness of buspirone and methylphenidate in treatment of attention deficit/hyperactivity disorder in children and adolescents. Iranian Journal of Psychiatry and Clinical Psychology. 2013;18(4):292-297.

[45] Barragan E, Breuer D, Dopfner M. Efficacy and Safety of Omega-3/6 Fatty Acids, Methylphenidate, and a Combined Treatment in Children With ADHD. J Atten Disord. 2014.

[46] Robb AS, Findling RL, Childress AC, Berry SA, Belden HW, Wigal SB. Efficacy, Safety, and Tolerability of a Novel Methylphenidate Extended-Release Oral Suspension (MEROS) in ADHD. J Atten Disord. 2014.

[47] Ghanizadeh A, Sayyari Z, Mohammadi MR. Effect of methylphenidate and folic Acid on ADHD symptoms and quality of life and aggression: a randomized double blind placebo controlled clinical trial. Iran J Psychiatry. 2013;8(3):108-112.

[48] Sobanski E, Dopfner M, Ose C, Fischer R. A non-interventional study of extended-release methylphenidate in the routine treatment of adolescents with ADHD: effectiveness, safety and adherence to treatment. Atten Defic Hyperact Disord. 2013;5(4):387-395.

[49] Garg J, Arun P, Chavan BS. Comparative short term efficacy and tolerability of methylphenidate and atomoxetine in attention deficit hyperactivity disorder. Indian Pediatr. 2014;51(7):550-554.

[50] Soutullo C, Banaschewski T, Lecendreux M, Johnson M, Zuddas A, Anderson C, et al. A post hoc comparison of the effects of lisdexamfetamine dimesylate and osmotic-release oral system methylphenidate on symptoms of attention-deficit hyperactivity disorder in children and adolescents. CNS Drugs. 2013;27(9):743-751.

[51] Evans C. The use of consensus methods and expert panels in pharmacoeconomic studies. Practical applications and methodological shortcomings. Pharmacoeconomics. 1997;12(2 Pt 1):121-129.

[52] Evans C, Crawford B. Expert judgement in pharmacoeconomic studies. Guidance and future use. Pharmacoeconomics. 2000;17(6):545-553.

[53] Danckaerts M, Sonuga-Barke EJ, Banaschewski T, Buitelaar J, Dopfner M, Hollis C, et al. The quality of life of children with attention deficit/hyperactivity disorder: a systematic review. Eur Child Adolesc Psychiatry. 2010;19(2):83-105.

[54] Graetz BW, Sawyer MG, Hazell PL, Arney F, Baghurst P. Validity of DSM-IVADHD subtypes in a nationally representative sample of Australian children and adolescents. J Am Acad Child Adolesc Psychiatry. 2001;40(12):1410-1417.

[55] Matza LS, Secnik K, Mannix S, Sallee FR. Parent-proxy EQ-5D ratings of children with attention-deficit hyperactivity disorder in the US and the UK. Pharmacoeconomics. 2005;23(8):777-790.

[56] Sawyer MG, Whaites L, Rey JM, Hazell PL, Graetz BW, Baghurst P. Health-related quality of life of children and adolescents with mental disorders. J Am Acad Child Adolesc Psychiatry. 2002;41(5):530-537.

[57] van der Kolk A, Bouwmans CA, Schawo SJ, Buitelaar JK, van Agthoven M, Hakkaart-van Roijen L. Association between quality of life and treatment response in children with attention Deficit Hyperactivity Disorder and their parents. J Ment Health Policy Econ. 2014;17(3):119-129.

[58] Brooks R, Rabin R, De Charro F. The measurement and valuation of health status using EQ-5D: a European perspective: evidence from the EuroQol BIOMED Research Programme. : Springer; 2003.

[59] Baldwin S, Gerard K. Caring at home for the children with mental handicaps. In: Baldwin S, Godfrey C, Propper C, editors. Quality of life: perspectives and policy London: Routledge; 1990.

[60] Parker G. Spouse Carers. Whose quality of life? In: Baldwin S, Godfrey C, Propper C, editors. The quality of life: perspectives and policies London: Routledge; 1990.

[61] Basu A, Meltzer D. Implications of spillover effects within the family for medical cost-effectiveness analysis. J Health Econ. 2005;24(4):751-773.

[62] De Meijer C, Brouwer W, Koopmanschap M, Van den Berg B, Van Exel J. The value of informal care--a further investigation of the feasibility of contingent valuation in informal caregivers. Health Econ. 2010;19(7):755-771.

[63] De Ridder A, De Graeve D. Healthcare use, social burden and costs of children with and without ADHD in Flanders, Belgium. Clin Drug Investig. 2006;26(2):75-90.

[64] Hakkaart L, Tan S, Bouwmans C. *Manual for costing research (in dutch).* Amstelveen: Healthcare insurance board; 2010.

[65] Ministry of Education, Culture and Science. Kerncijfers 2007-2011 onderwijs, cultuur, wetenschap. 2012.

[66] IMS Health BV. Patient Level Study Concerta. 2008.

[67] Zorginstituut Nederland. Medicijnkosten. 2014. http://www.medicijnkosten.nl. Accessed 12/18, 2014.

[68] Barkley RA, Fischer M, Smallish L, Fletcher K. Young adult follow-up of hyperactive children: antisocial activities and drug use. J Child Psychol Psychiatry. 2004;45(2):195-211.

[69] Sourander A, Elonheimo H, Niemela S, Nuutila AM, Helenius H, Sillanmaki L, et al. Childhood predictors of male criminality: a prospective population-based follow-up study from age 8 to late adolescence. J Am Acad Child Adolesc Psychiatry. 2006;45(5):578-586.

[70] Mannuzza S, Klein RG, Bessler A, Malloy P, LaPadula M. Adult psychiatric status of hyperactive boys grown up. Am J Psychiatry. 1998;155(4):493-498.

[71] Vitelli R. Prevalence of Childhood Conduct and Attention-Deficit Hyperactivity Disorders in Adult Maximum-Security Inmates. International Journal of Offender Therapy and Comparative Criminology. 1996;40(4):263-271.

[72] Lichtenstein P, Halldner L, Zetterqvist J, Sjolander A, Serlachius E, Fazel S, et al. Medication for attention deficit-hyperactivity disorder and criminality. N Engl J Med. 2012;367(21):2006-2014.

[73] Loe IM, Feldman HM. Academic and educational outcomes of children with ADHD. J Pediatr Psychol. 2007;32(6):643-654.

[74] Mannuzza S, Klein RG, Bessler A, Malloy P, LaPadula M. Adult outcome of hyperactive boys. Educational achievement, occupational rank, and psychiatric status. Arch Gen Psychiatry. 1993;50(7):565-576.

[75] Teeter P. Interventions for ADHD: Treatment in Developmental context. New York: Guilford Press; 1998.

[76] Biederman J, Petty CR, Fried R, Kaiser R, Dolan CR, Schoenfeld S, et al. Educational and occupational underattainment in adults with attention-deficit/hyperactivity disorder: a controlled study. J Clin Psychiatry. 2008;69(8):1217-1222.

[77] Biederman J, Faraone SV. The effects of attention-deficit/hyperactivity disorder on employment and household income. MedGenMed. 2006;8(3):12.

[78] Klein RG, Mannuzza S, Olazagasti MA, Roizen E, Hutchison JA, Lashua EC, et al. Clinical and Functional Outcome of Childhood Attention-Deficit/Hyperactivity Disorder 33 Years Later. Arch Gen Psychiatry. 2012:1-9.

[79] Swanson J. Compliance with stimulants for attention-deficit/hyperactivity disorder: issues and approaches for improvement. CNS Drugs. 2003;17(2):117-131.

[80] Fenwick E, Byford S. A guide to cost-effectiveness acceptability curves. Br J Psychiatry. 2005;187:106-108.

[81] Fenwick E, O'Brien BJ, Briggs A. Cost-effectiveness acceptability curves--facts, fallacies and frequently asked questions. Health Econ. 2004;13(5):405-415.

[82] Petrou S. Methodological issues raised by preference-based approaches to measuring the health status of children. Health Econ. 2003;12(8):697-702.

[83] Vogels T, Verrips GH, Verloove-Vanhorick SP, Fekkes M, Kamphuis RP, Koopman HM, et al. Measuring health-related quality of life in children: the development of the TACQOL parent form. Qual Life Res. 1998;7(5):457-465.

[84] Juniper EF, Guyatt GH, Feeny DH, Griffith LE, Ferrie PJ. Minimum skills required by children to complete health-related quality of life instruments for asthma: comparison of measurement properties. Eur Respir J. 1997;10(10):2285-2294.

[85] Sprangers MA, Aaronson NK. The role of health care providers and significant others in evaluating the quality of life of patients with chronic disease: a review. J Clin Epidemiol. 1992;45(7):743-760.

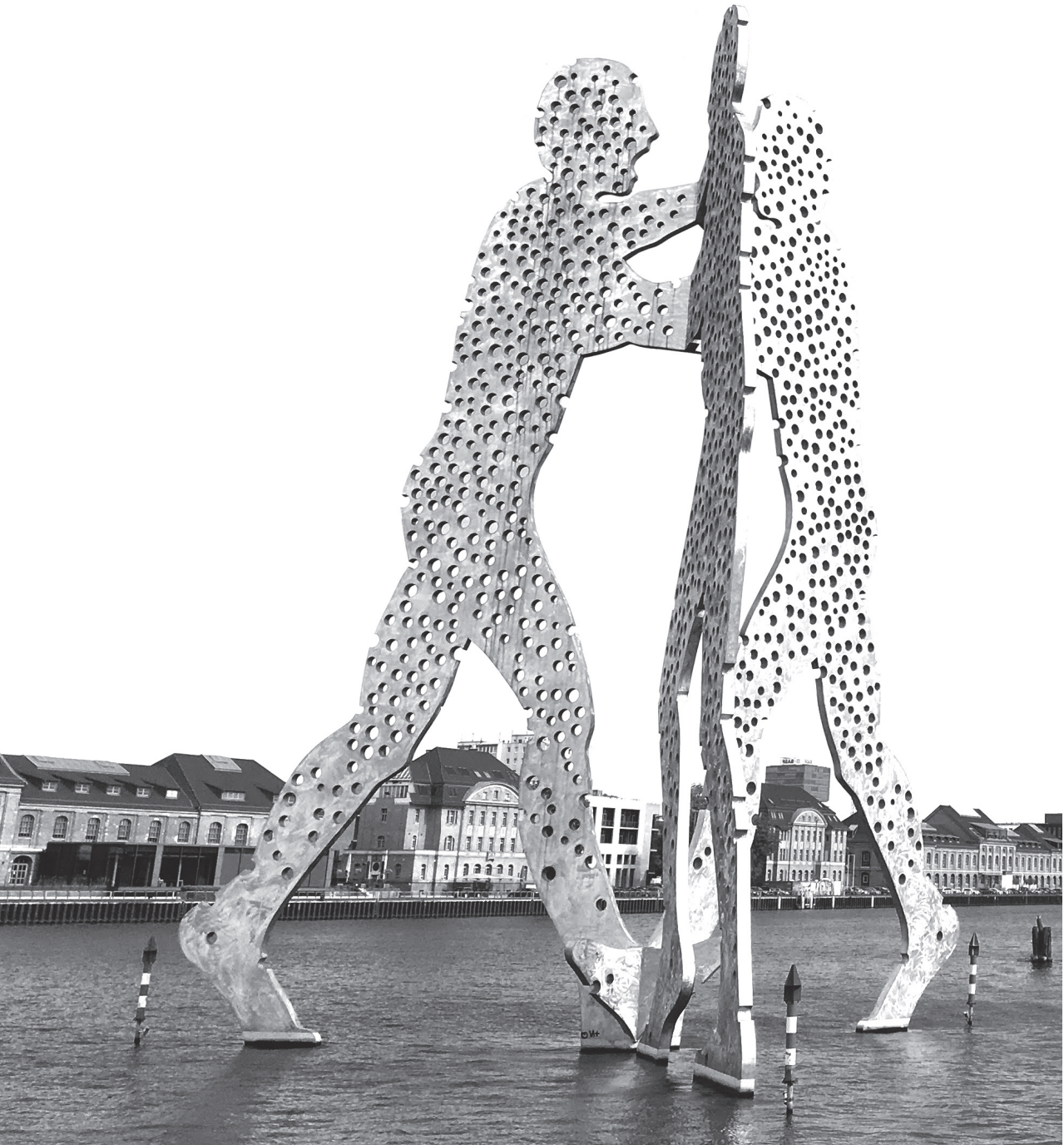
[86] Glaser AW, Davies K, Walker D, Brazier D. Influence of proxy respondents and mode of administration on health status assessment following central nervous system tumours in childhood. Qual Life Res. 1997;6(1):43-53.

[87] Romeo R, Byford S, Knapp M. Annotation: Economic evaluations of child and adolescent mental health interventions: a systematic review. J Child Psychol Psychiatry. 2005;46(9):919-930.

[88] College tarieven gezondheidszorg. Tariffs AWBZ-institutions 2005. 2013. http://www.ctg.zaio.nl. , 2013.

[89] Dutch Society of Remedial Teachers. Standard prices. 2013. http://www.lbrt.nl. , 2013.

[90] Health Care Insurance Board. Personal budget. 2013. http://www.cvz.nl/cijfers_en_publicaties/brochures_en_folders/zorginkopen_pgb.asp. , 2013.



Chapter 3

The cost-effectiveness of family/family-based therapy for treatment of externalizing disorders, substance use disorders and delinquency: A Systematic review

Based on Goorden, M.*, Schawo, S.*, Bouwmans-Frijters, C., van der Schee, E., Hendriiks, V., Hakkaart-van Roijen, L.

BMC Psychiatry; 2016:16(1), 237.

Introduction

Family therapy and family-based treatment is considered an evidence-based practice treatment for children and adolescents with externalizing disorders, symptoms of delinquency and/or substance use disorder [1, 2]. Familial and extra-familial systems are known to influence the individual [3-7], and therefore family/family-based therapy is not only aimed at the individual youth but also at systems surrounding the individual. For instance, delinquency and substance abuse in adolescents have been shown to be influenced by family factors, like parenting style and attachment [3-7]. In addition, a recent review indicated that problems within the extra-familial system, like delinquent peers, problems with bonding at school and in the neighborhood are risk factors for delinquency and problem drinking [7]. As the individual, familial and extrafamilial systems are interconnected, family/family-based therapy not only positively affects the adolescent but also the family (family cohesion) and the extra-familial systems [8].

For the purpose of the present paper, family therapy and family-based treatment is broadly defined as treatments in which primarily family members and/or members of the families' wider networks are involved in the treatment process of resolving problems for young people [9] as opposed to treatments that mainly or solely focus on the individual youth, or treatments that do not focus on youths' problem behavior, like marital therapy.

Well-known forms of family/family-based treatments are Multisystemic therapy (MST) [10], Functional Family Therapy (FFT) [11] and Multidimensional Family therapy (MDFT) [12]. Although there is a large overlap between these types of therapies, there are also some differences [13]. For instance, in FFT and MST there is more focus on antisocial behavior. However, the degree of severity of the disorder is often higher in MST compared to FFT. More details of these differences are described in Appendix 3.1. Recently, Von Sydow et al. [1] systematically reviewed studies on the effectiveness of family/family-based therapy for the treatment of children and adolescents who have externalizing disorders. Their study included disorders like substance abuse, attention deficit hyperactivity disorder, conduct disorder and symptoms of delinquency. They concluded that there is sound evidence that family/family-based therapy is effective with particularly large effect sizes for delinquency and substance abuse measures. However, in the meta analyses that were included in Von Sydow's systematic review, more cautious conclusions regarding the effectiveness of

systemic therapy were drawn. Current health care policy in the Netherlands and elsewhere places emphasis on the provision of effective mental health services in a cost effective way. Family/ family-based interventions are intensive as they consist of a relatively high number of sessions per week and subsequently are relatively expensive [14-16]. Therefore, there is a need for economic evaluations to assess whether additional effects gained through family/family-based therapy in comparison to alternative treatments – if observed – justify the additional costs. Morgan et al. [17] described eight studies, analyzing the cost-effectiveness of family-based treatments for substance abusing adults and adolescents and concluded that some of these treatments could be considered as cost-effective. However, family based therapies like marital therapy, were also included in this study. In addition, the literature search in this study was not systematically conducted and was only considering patients with substance use disorders. To our knowledge, no systematic review of economic evaluations of family/family-based therapy in externalizing, delinquent or substance-abusing adolescents has yet been performed. Hence, this paper presents a systematic review of economic evaluations of systemic interventions in adolescents with externalizing disorders, substance abuse or delinquency.

The aim of the present study was to assess the evidence on cost-effectiveness of family/family-based therapy for adolescents with externalizing disorders, substance use disorders or delinquency, and to evaluate the quality of the existing studies, and the generalizability of the study findings.

Methods

The review was performed according to the Cochrane handbook for systematic reviews of interventions [18] and adopted the Preferred Reporting for Systematic reviews and Meta-Analyses (PRISMA) statement [19].

Search strategy

A systematic literature search was performed in Pubmed, ERIC, Psycinfo and Cochrane reviews (including economic trials and clinical trials). These different search engines were used because of their high quality, coverage of large databases and their focus on economic trials. Search terms encompassed the different types of systemic therapy (Functional Family Therapy, Multidimensional Family therapy, Multidimensional Foster Care, Multisystemic Therapy, Family Behavior Therapy and Brief Strategic Therapy) but also more general classifications (systemic therapy, substance abuse treatment, family based therapy, Family

based intervention, Family system intervention, Family intervention program). These terms were searched for in title and abstract and were then combined with terms referring to economic evaluations searched for in title and abstract or a Medical Subject Headings (MeSH) term (economic evaluation, cost-effectiveness, cost-utility, cost benefit, cost analysis, cost measure) and in the title (costs). Costs were searched for only in the title, and not in the abstract, because the latter resulted in many irrelevant studies. This search term was included as we noticed that although in some studies both costs and effects were evaluated, the main focus of these studies was to evaluate the costs and a smaller part was referring to the effects. Consequently, when only terms referring to both the costs and effects were included, these studies would have been missed. The search term "Economic modeling" was not explicitly incorporated into the search strategy as the modeling should be part of a cost-effectiveness, cost utility, cost benefit or cost analysis (corresponding with our aim). Abbreviations were also included. To improve our search, MeSH terms were used, see Appendix 3.2 for more details.

Selection strategy

In Fig. 3.1 the selection criteria are described and numbered. The criteria were applied to the studies in chronological order and when a study was excluded based on a criterion the number as shown in Fig. 3.1 was noted. We considered studies from January 1990 until January 2016. The selected study types were clinical/randomized controlled trials (RCT), reviews, systematic reviews and meta-analyses. The treatment needed to consist of a family/family-based intervention, targeted at adolescents (10–20 years old) with a substance use disorder, externalizing disorder or delinquent behavior. The method needed to be a cost or cost- effectiveness/benefit/ utility analysis. When studies were assessed for eligibility based on their abstracts and it was likely that they only contained cost-outcomes and no effect-outcomes, they were also included. To determine the eligibility of the full text articles, the same selection criteria were used, except that accessibility of the study was a requirement (full text available) and studies that only contained costs-outcomes and no effect-outcomes were excluded. The selection of the articles was performed by two researchers independently. Differences in selections were discussed until consensus was reached.

Data extraction and risk of bias

The quality of the studies was assessed with the British Medical Journal Checklist for authors and peer reviewers of economic submissions [20] and the Consensus

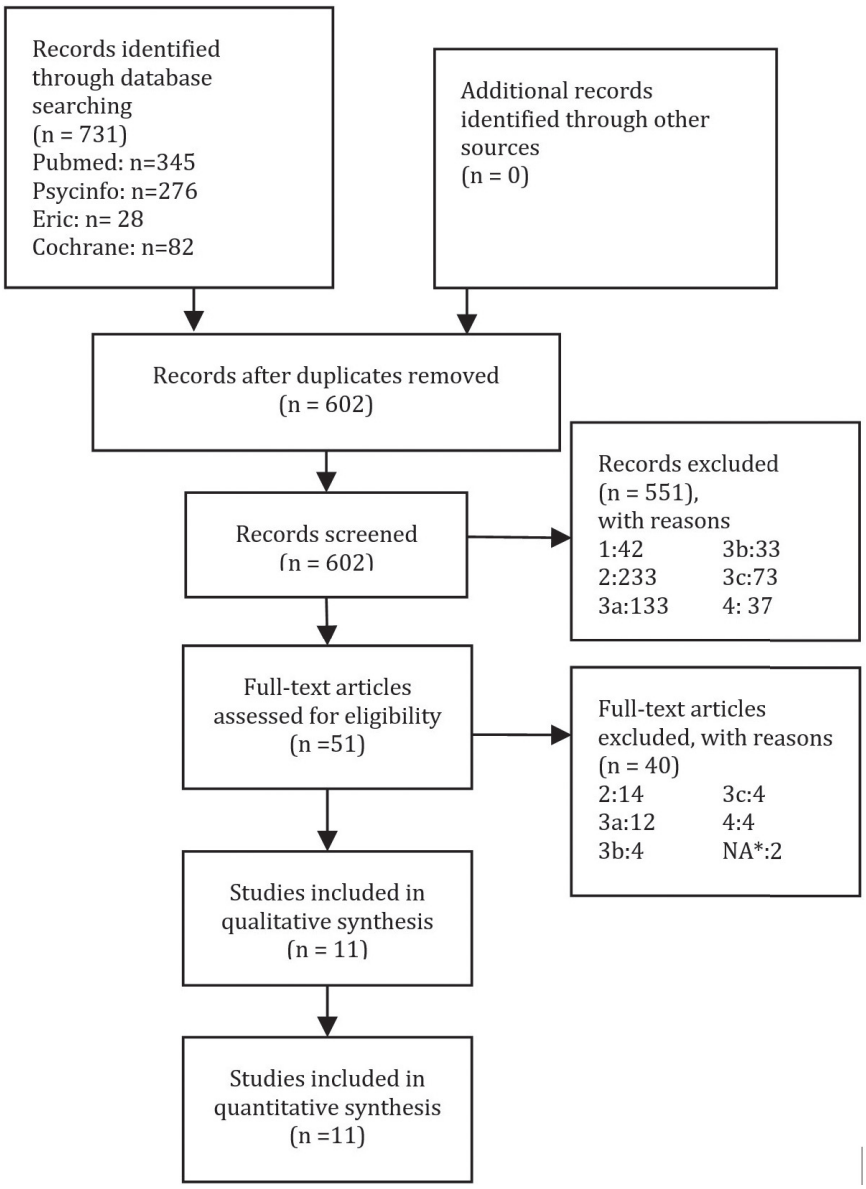
on Health Economic Criteria (CHEC) list for assessment of methodological quality of economic evaluations [21] as recommended by the Cochrane reviews handbook [18]. We also consulted the critical appraisal of the studies by the NHS Economic Evaluation Database (NHS EED) structured abstract [195]. This is a database from Cochrane library consisting of structured abstracts of economic evaluations of health care interventions. Full economic evaluations were identified from a variety of sources and assessed according to a set of quality criteria. Subsequently, detailed structured abstracts were produced. In addition to the checklists, information about the economic perspective of the study (health care, societal etc.), design, country, follow-up, type of disorder, sample size, study dropout, age, gender, type, duration and intensity of intervention, time horizon, currency and price year, key features of sensitivity analyses and the included cost types were collected for the economic evaluation described in the studies. In accordance with the suggestions in the Cochrane handbook [18] five different biases of the individual studies were addressed: selection bias, performance bias, detection bias, attrition bias and reporting bias [18]. They were respectively addressed by assessing if patients were properly balanced at baseline, patients and therapists were blinded, outcome assessors were blinded, the amount of dropout in the studies and by reading the protocols of the studies.

Results

A total of 731 articles met the search criteria. After removal of duplicates and a first selection based on the abstracts, 51 studies matched the inclusion criteria. After assessment for eligibility, 11 studies were selected (see Fig. 3.1).

Figure 1 | PRISMA flow diagram [19]

Selection/eligibility criteria		
1.	Year of publication	After 1990
2.	Publication:	Systematic Review, randomized/clinical trial, meta-analysis
3.	a Treatment:	family/family-based
	b Population:	Adolescents, 10-20 years of age
	c Outcome:	Cost-effectiveness/benefit/utility, costs
4.	Disease/symptoms:	delinquency, substance use disorders, externalizing



*NA=Not available; Two studies were not available

Characteristics of the studies

An overview of the characteristics of the studies, participants and the interventions is shown in Table 3.1. Ten of the eleven selected studies were published between 2003 and 2015 [22-31] and one study was published in 1996 [207]. Eight of the studies originated from the United States (USA) [22-24, 27, 29-32].

Remaining studies were initiated in Sweden [26] ,England [28] and Mexico [25]. All studies were (based upon) randomized controlled trials. Two pairs of studies [22, 24, 27, 29] were each based on one sample. Most of the studies compared a family/family-based intervention with care as usual [23, 26, 28, 30-32]. MST was the most researched intervention as it was investigated in eight studies [23, 24, 26-28, 30-32]. In the Study of Borduin et al. [31] Multisystemic Therapy for Problem Sexual Behavior (MST-PSB) was investigated. MST-PSB is an adaptation to MST aimed at the treatment of juvenile sexual offenders. A description of the (non- family/family-based) comparator interventions is shown in Table 3.2. The mean number of sessions of the family/family-based interventions was between 1 and 3 times a week and the mean duration of treatment was between 12 and 31 weeks. The average follow-up time was between 6 and 300 months (25 years); only four studies followed patients for more than 1 year [26, 28, 30, 31]. Two studies were outliers in respect to the time horizon they used (8 years and 25 years) [30,31].

Six studies were aimed at adolescents with substance use disorder [22, 24, 25, 27, 29, 32], one study investigated adolescents with a conduct disorder [26], one study adolescents at risk for continuing criminal activity [26], one study adolescents who had experienced a psychiatric crisis [23], another study adolescents who were serious juvenile offenders [30] and one study aimed to investigate juvenile sexual offenders [31]. The average sample size of the 9 studies (with separate samples) was 178 (SD = 163) with a variation between 48 and 600 patients. Follow-up attrition, when registered, was low (not more than 30%). Average age at baseline was 15 (Standard Deviation (SD) = 1) years and between 61 and 96% of the individuals were males. Types of economic analyses included cost-effectiveness analyses [23, 25, 27, 29], cost-benefit analyses [22, 26, 30, 31] and cost offset analyses [28]. The difference between a cost-offset and a cost-benefit analysis is often not well-explained. A cost-offset analysis compares the monetary value of resource use with the monetary value of costs reduced by the intervention (usually health care costs). In contrast to a cost-benefit analysis which also focuses on other outcomes that are translated in monetary outcomes (like translating number of life years gained to a monetary value). In reality, cost-offset analysis is a partial cost-benefit analysis because it compares the cost of a program with the monetary value of a single outcome (i.e., avoided future health care costs). In two studies, the economic evaluation was not explicitly classified [24, 32].

Table 1 | Features of the studies, participants and the interventions

Study	Features study			Features participants								Features intervention		
	Country	Follow-up (months)	Design	Disorder	Sample Size		Completed study		Age		Sex (% male)		Intervention	Number of sessions per week	Treatment, duration (weeks)
					I	C	I	C	I	C	I	C			
Schoenwald et al.,1996 [207]	USA	6 AT	RCT	SUD	59	59	NS		16		79		CAU	2–3[b]	18–19
French et al., 2003 [197]	USA	12a	RCT	SUD	102		564		16		81		Trial 1:	0–1	6–7
					96				16		86		MET/CBT5	0–1	12–14
					102				16		84		MET/CBT12	1–2	12–14
													FSN		
					100				16		79		Trial 2:	0–1	6–7
					100				16		80		MET/CBT5	1–2	12–14
					100				16		85		ACRA	1–2	12–14
													MDFT		
Sheidow et al.,2004 [198]	USA	12 AT	RCT	PC	115		NS		13		67		MST CAU	NS	16
Dennis et al., 2004 [204]	USA	12	RCT	SUD	102		564		16		81		Trial 1:	0–1	6–7
					96				16		86		MET/	0–1	12–14
					102				16		84		CBT5	1–2	12–14
													MET/		
					100				16		79		CBT12	0–1	6–7
					100				16		80		FSN	1–2	12–14
					100				16		85		Trial 2:	1–2	12–14
													MET/		
													CBT5		
													ACRA		
													MDFT		

Table 1 | Continued

Study	Features study			Features participants								Features intervention			
	Country	Follow-up (months)	Design	Disorder	Sample Size		Completed study		Age		Sex (% male)		Intervention		Number of sessions per week	Treatment, duration (weeks)
					I	C	I	C	I	C	I	C				
McCollister et al.,2009 [199]	USA	12	RCT	SUD	38 38 43	42	NS		15 15 15	15	84 84 84	81	DC DC + MST DC + MST + CM	FC	NS[1]	NS[1]
French et al.,2008 [200]	MEX	7	RCT	SUD	30 29 31	30	114		16 16 16	16	80 76 84	83	FFT Joint CBT	group	NS[1]	NS[1]
Olsson, 2010 [201]	SW	24	RCT	CD	79	77	NS		15		61		MST	CAU	NS	12–20
Sheidow et al.,2012 [202]	USA	12	RCT	SUD	38 38 43	42	29 29 37	33	15		83		DC DC + MST DC + MST + CM	FC	NS[1]	NS[1]
Cary et al., 2013 [203]	ENG	30	RCT	DEL	56	52	46	45	15	15	83	82	MST+ CAU	CAU	3	20
Dopp et al.(2014) [205]	USA	300	RCT	DEL	92	84	70	56	15		69		MST	CAU	3–4	21
Borduin et al.(2015) [206]	USA	107	RCT	DEL	24	24	24	22	14		96		MST-PSB	CAU	3	31

Legend: I =intervention, C=comparator, NS=not stated,NS[1]=reference to non-accessible article, NA=not applicable, USA=United States of America, SW=Sweden, ENG= England, MEX=Mexico,SUD=substance use disorder, CD=conduct disorder, PC=psychiatric crisis, MST=multisystemic therapy, Joint=combination of individual and family therapy, group=skill-focused psycho-education group intervention, IT= individual treatment, MST-PSB=MST for Problem Sexual Behavior, CAU=care as usual, FSN=family support network, MDFT=multidimensional family treatment, MET/CBT12=motivational enhancement treatment/cognitive behavior therapy, 12 sessions;MET/CBT5=motivational enhancement treatment/cognitive behavior therapy, 5 sessions, ACRA= adolescent community reinforcement approach, DC=drug court with community services, DC +MST=drug court with multisystemic therapy, DC +MST + CM=drug court with mst and enhanced with a contingency management programs, FFT=functional family therapy, FC=family court with community services, [a] Cost data was only collected only during 3–9 months, [b] The intensity of the treatment was between 2 and 3 times a week; AT=after treatment

Table 2| Descriptions of comparator interventions

FSN	Cognitive behavioral sessions and motivation treatment in combination with a family component
MET/CBT5	Motivational component and a cognitive behavioral component, to enhance motivation to change drug abuse and to grow the skills to maintain and regulate abstinence
MET/CBT12	MET/CBT5+ 7 sessions of CBT are added to the therapy.
FC	Family court treatment with community services/ Appearance court 2 times a year/ outpatient alcohol and drug abuse service from the local center of the state's substance abuse commission
DC	Drug court treatment with community services/ Appearance court 1 time a week/ outpatient alcohol and drug abuse service from the local center of the state's substance abuse commission and monitoring drug abuse
CM	Frequent in home screens for drug use, voucher system contingent on clean screens, and drug refusal training.
ACRA	Identifying reinforces that are incompatible with the drug use and to strengthen those
CAU	Sheidow et al. [198] admission to a psychiatric unit and aftercare Schoenwald et al. [207] outpatient substance abuse services Olsson et al. [201] Not described Cary at al. [203] Youth Offending Team (YOT) Dopp et al. [205] Individual Therapy (IT) Borduin et al. [206] Cognitive behavioral group therapy and individual services (from local juvenile court)

FSN family support network, MET/CBT5 motivational enhancement treatment/ cognitive behavior therapy, 5 sessions, MET/CBT12 motivational enhancement treatment/cognitive behavior therapy, 12 sessions; ACRA adolescent community reinforcement approach, FC family court with community services, DC drug court with community services, CM contingency management programs, CAU care as usual

Outcomes of the studies

Details of the interventions and outcomes of our analyses are described in Tables 3.3 and 3.4. Costs were indexed until 2014.

Table 3| Studies that reported substance use disorder

Studies considering costs and effects of substance abuse			
Dennis (2004)	**Costs intervention and comparators (per episode of care per patient)(MET/CBT 5, MET/CBT 12, FSN, ACRA, MDFT)** In trial 1 MET/CBT 5,MET/CBT 12 and FSN were compared. In trial 2 MET/CBT 5, ACRA and MDFT were compared. Costs were collected with a program (DATCAP), which yields estimates such as the total annual opportunity cost of treatment and the labor cost per client.		**Difference cost** The differences in costs were not shown in this study. However, it was showed that the differences were significant.
	MET/CBT 5 (trial 1): €1,226	MET/CBT 5 (trial 2): €1,716	
	MET/CBT 12 (trial 1): €1,305	ACRA (trial 2): €1,551	
	FSN (trial 1): €3,576	MDFT (trial 2): €2,205	
	Effects intervention and comparators (per patient) (MET/CBT 5, MET/CBT 12, FSN, ACRA, MDFT)		**Difference effects**
	Met CBT 5 (trial 1) Days of abstinence: 269 Recovery*: 28%	Met CBT 5 (trial 2) Days of abstinence: 251 Recovery: 23%	The differences in effects were not shown in this study However it was showed that the difference was not significant.
	Met CBT 12 (trial 1) Days of abstinence: 256 Recovery: 17%	ACRA (trial 2) Days of abstinence: 265 Recovery: 34%	
	MET FSN (trial 1) Days of abstinence: 260 Recovery: 22%	MDFT (trial 2) Days of abstinence: 257 Recovery: 19%	
	*Recovery is defined as having no use or abuse dependence problems and living in the community		
Results	**Cost per day of abstinence:** Met CBT5 (trial 1): €541 Met CBT 12: €677 Met FSN: €1,667		**Cost per days of abstinence:** MET/CBT5 (trial 2): €991 ACRA: €729 MDFT: €1,143
	Costs per person in recovery Met CBT5 (trial 1): €4,360 Met CBT 12: €41,172 Met FSN: €16,651		**Costs per person in recovery** MET/CBT5 (trial 2): € 7,337 ACRA: €4,913 MDFT: €12,970

Studies considering costs and effects of substance abuse			
French (2008)	**Costs intervention per patient (FFT, Joint and CBT)** FFT: Treatment costs: €1,817 Joint: treatment costs: €2,847 CBT: Treatment costs: €1.439	**Costs comparator per patient (Group)** Group: Treatment costs: €990	**Difference costs** The difference in costs were not showed in this study
	Effects intervention per patient (FFT, Joint and CBT) FFT: % days marijuana use 4 months: 25.3 % of days marijuana use 7 months: 39.8 YSR delinquency score 4 months: 8.2 YSR delinquency score 7 months: 9.2 Joint % of days of marijuana use 4 months: 38.1 marijuana use 7 months: 35.4 YSR delinquency score 4 months: 9.1 YSR delinquency score 7 months: 8.5 CBT %of days marijuana use 4 months: 50.6 %of days marijuana use 7 months:51.8 YSR delinquency score 4 months: 10.2 YSR delinquency score 7 months:10.4	**Effects comparator per patient (Group)** Group % of days of marijuana use 4 months: 54.8 marijuana use 7 months: 40.7 YSR delinquency score 4 months: 9.5 YSR delinquency score 7 months: 9.4	**Difference effects with regression model:** FFT versus group: % days marijuana use 4 months: 20.11* % days marijuana use 7 months: 4.87 YSR delinquency score 4 months: -0.60 YSR delinquency score 7 months: 0.15 CBT versus group % days marijuana use 4 months: 4.76 % days marijuana use after 7 months: 18.27 YSR delinquency score 4 months: 0.38 YSR delinquency score 7 months: 0.42 Joint versus group % days marijuana use after 4 months: -14.86 % days marijuana use after 7 months: -2.00 YSR delinquency score 4 months: -0.50 YSR delinquency score 7 months: -1.50
Results	Group therapy was most cost-effective, none of the other therapies were significantly different in effect compared to group therapy. So the intervention with the lowest costs was considered to be most cost-effective.		

Table 3 | Continued

Studies considering costs and effects of substance abuse			
Sheidow (2012)	**Costs Intervention (DC, DC+MST, DC+MST+CM)** Treatment costs DC: €9,083 DC+MST: €12,369 DC+MST+CM: €12,859	**Costs comparator (FC)** Treatment costs FC: €3,679	The difference in costs were not shown in this study
	Effects intervention (DC, DC+MST, DC+MST+CM)	**Effects comparator (FC)**	**Difference effects**:
	DC Marijuana use (days): -16.65 Polydrug use (days): 1.41 Alcohol use (days): 0.49 Heavy alcohol use (days): 0.86 SRD status offenses (incidents): -7.24 SRD Theft (incidents): -3.28 SRD crimes against persons (incidents): -2.69	Marijuana use (days): -15,43 Polydrug use (days): 2.27 Alcohol use (days): 2.97 Heavy alcohol use (days): 0.76 SRD status offenses (incidents): 9.22 SRD Theft (incidents): -5.54 SRD crimes against persons (incidents): 0.49	The difference in effects were not showed in this study
	DC+MST Marijuana use (days): -30,17 Polydrug use (days): :-1.11 Alcohol use (days): 0.27 Heavy alcohol use (days): -0.45 SRD status offenses (incidents): -11.11 SRD Theft (incidents): -2.79 SRD crimes against persons (incidents): -3.90		
	DC+MST+CM Marijuana use (days): -27.86 Polydrug use (days): -6.76 Alcohol use (days): -7.56 Heavy alcohol use (days): -4.13 SRD status offenses (incidents): -10.38 SRD Theft (incidents): -3.19 SRD crimes against persons (incidents): -2.4		
Results	ACERS (Average cost-effectiveness ratios) were calculated; average costs/ difference between mean incidents before and after treatment(negative means inefficient)		

Table 3 | Continued

	FC	DC	DC+MST	DC+MST+CM
Marijuana use:	€238	€545	€410	€461
	(215-262)	(474-617)	(377-442)	(434-488)
Polydrug use:	€-1,619	€-6,425	€11,209	€1,912
	(-8,839-5,601)	(-27,541-14,692)	(-3,757-26,175)	(1,624-2,182)
Alcohol use:	€-,1,239	€-18,814	€-44,838	€1,699
	(-6,546-5,601)	(-42,034-4,405)	(-61,014-28,662)	(1,486-1,912)
Heavy alcohol use:	€-4,857	€-10,535	€27,592	€3,109
	(-10,632-918)	(-28,804-7,733)	(-14,636-69,821)	(1,708-4,511)
SRD status offenses:	€-400	€1,254	€1,114	€1,239
	(-1,206-398)	(1,132-1,376)	(907-1,321)	(1,009-1,496)
SRD theft:	€663	€2,773	€4,428	€4,032
	(428-899)	(-2.441-7,987)	(-1,224-10,081)	(1,204-6,859)
SRD crimes against persons:	€-7,588	€3,377	€3,175	€5,346
	(-10,667--4,510)	(2,976-3,777)	(236-6,123)	(4,723-5,968)

Table 3 | Continued

Studies considering costs and effects of substance abuse				
Schoenwald (1996)	**Costs interventions(MST)** Mental health outpatient (total): €4,242 Mental health day treatment (total): €5,423 Mental health residential treatment (total): €6,899 Psychiatric inpatient (total): €15,752 Psychiatric emergency room (total): €1,150 Substance abuse outpatient (total): €2,001 Substance abuse residential treatment (total): €3,450 Substance abuse inpatient (total): €16,098 Marine Institute day treatment (total): €18,926 Marine Institute residential treatment (total): €3,036 Treatment costs: €266,516	**Costs comparator (CAU)** Mental health outpatient (total): €19,075 Mental health day treatment (total): €1,118 Mental health residential treatment (total): €0 Psychiatric inpatient (total): €18,513 Psychiatric emergency room (total): €3,450 Substance abuse outpatient (total): €20,272 Substance abuse residential treatment (total): €43,695 Substance abuse inpatient (total): €93,771 Marine Institute day treatment (total): €28,618 Marine Institute residential treatment (total): €0	**Benefits interventions** Incarceration days: €65,427	**Benefits CAU** Incarceration days: €120,851
Results	MST: Total costs with incarceration=€408,919 and the total costs with incarceration per youth=€6,930 CAU: Total costs with incarceration=€335,845 and the costs per youth=€5,693. Difference in total between groups =€1,019			

Table 3 | Continued

Studies considering costs and effects of substance abuse				
French (2003)	**Costs interventions (MET/CBT 5, MET/CBT 12, FSN, ACRA, MDFT)**		**Benefits interventions (MET/CBT 5, MET/CBT 12, FSN, ACRA, MDFT)**	
	Treatment costs were measured		**Health service utilization;** Outpatient clinic/doctor's office visit Days bothered by health/medical problem **Substance-abuse treatment** utilization; Days in detoxification program; Day in inpatient treatment program; Day in long-term residential program; Intensive outpatient program visits; Regular outpatient program visits **Education and employment;** Days missed at school or training; Personal income; Days stressful for parents; Day missed of work or school by parent **Criminal activity;** Arrests; Day on probation; Days on parole; Days in prison/jail; Days in juvenile detention	
	Incremental arm:	Alternative arm:	**Incremental arm:**	**Alternative arm:**
	MET/CBT5:€1,226 MET/CBT12: €1,305 FSN: €3,576	MET/CBT5: €1,716 ACRA: €1,551 MDFT: €2,216	MET/CBT5 Baseline €2,553 3 months: €2,133 6 months: €1,671 9 months: €945 12 months: €1,217 MET/CBT12 Baseline: €2,179 3 months: €2,433 6 months: €828 9 months: €1,431 12 months: €687 FSN: Baseline: €2,552 3 months: €4,525 6 months: €1,783 9 months: €1,205 12 months: €1,726	MET/CBT5 Baseline €2,694 3 months: €3,587 6 months: €2,213 9 months: €2,275 12 months: €1,907 ACRA Baseline: €2,506 3 months: €3,691 6 months: €1,748 9 months: €3,113 12 months: €3,237 MDFT: Baseline: €2,019 3 months: €3,938 6 months: €1,467 9 months: €2,573 12 months: €2,098

Table 3 | Continued

Results	Net economic benefits (benefits+costs) relative to baseline: 3 different models were administered; Model 1: only time dummies for each of the follow-up periods (as treatment conditions were not included, we did not show the results. Model 2: time dummies and indicator variables for treatment condition. Model 3: time and treatment variables with an indicator variable for site. The last specification added numerous demographic and environmental controls.
	MET/CBT12: Model 2: €-198 (349) Model 3: €-171 (346) Model 4: €-340 (334)
	Acra: Model 2: €369 (436) Model 3: €530 (430) Model 4: €554 (405)
	FSN: Model 2: €607* (343) Model 3: €653 (340) Model 4: €250 (333) *p<0.1
	MDFT Model 2: -€61 (441) Model 3: €128(436) Model 4: €100 (530)

Table 3 | Continued

Studies considering costs and effects of substance abuse				
McCollister (2009)	**Costs interventions (DC, DC/MST, DC)**	**Costs comparators (FC)**	**Benefits interventions (DC, DC/MST, DC)**	**Benefits comparators (FC)**
	Treatment costs	Treatment costs	Criminal activity costs according to Self-reported criminal activity (SRD):	Self-reported criminal activity (SRD):
	DC: €8,156	FC: €3,304	DC: €28.601 (94.314)	FC: €206.045 (545.581)
	DC/MST: €11,547		DC/MST: €65.640 (240.559)	
	DC/MST/CM: €11,547		DC/MST/CM: €80.222 (336.461)	
Results	After 12 months, total costs relative to FC with multivariate model (intervention costs not incorporated): DC: €-124,877 (-84,107) DC/MST: €-117,918 (-82,570) DC/MST/CM: €140,274 (179,066) * All DC conditions generated reduction in crime costs, greater than average costs of treatment.			

Currency and price year: Sheidow (2004).USD, 1997; Dennis (2004).USD, 1999; French (2008). USD, 1998; Sheidow (2012).USD 2004. When a price year was not stated it was estimated by taking the mean year of the study duration or when not available subtracting 1 from the year of publication of the study. MST=multisystemic therapy; Joint=Combination of individual and family therapy; group=skill-focused psycho-education group intervention; CAU=Care As Usual; FSN= Family support network; MDFT= multidimensional family treatment; MET/CBT12: Motivational enhancement treatment/ cognitive behavior therapy, 12 sessions; MET/CBT5= Motivational enhancement treatment/ cognitive behavior therapy, 5 sessions ; Acra= Adolescent community reinforcement approach; DC=Drug Court with community services; DC+MST=Drug court with Multisystemic therapy; DC+MST+CM=Drug court with MST and enhanced with a contingency management programs; FFT= functional family therapy; FC= Family court with community services

Table 4| Studies considering externalizing disorders and delinquency

Studies considering both costs and effects				
Sheidow (2004)	**Costs intervention (MST)** Medicaid (government insurance program)costs (inpatient, Outpatient, Pharmacy, other costs), Other treatment costs paid for by study <u>MST</u> Medicaid costs: 0-end treatment (4 months): €9,311 (±7,755) Medicaid costs: End treatment-12 months: €13,237 (±15,144) Other treatment costs paid for by study: €11,617	**Costs comparator (CAU)** Medicaid (government insurance program)costs (inpatient, Outpatient, Pharmacy, other costs), Other treatment costs paid for by study <u>CAU</u> Medicaid costs: 0-end treatment (4 months): €13,255 (±5,762) Medicaid costs: End treatment-12 months: €15,207 (±18,485) Other treatment costs paid for by study: €0	**Difference costs ($Costs_{CAU}$-$Costs_{MST}$) (after risk adjusted model):** 0-end treatment (total costs): -€1,828 End treatment- 12 months post-treatment (total costs): -€452 (SE=14)	
	Effects intervention CBCL: Externalizing scores, internalizing scores: GSI: Global severity index are measures The main effects were not showed in this study but only differences over time were presented.	**Effects comparator** CBCL: Externalizing scores, internalizing scores: GSI: Global severity index The main effects were not showed in this study but only differences over time were presented.	**Difference effects ($Effects_{CAU}$-$Effects_{MST}$) (after risk adjusted model):** 0-end treatment: end treatment- 12 months post-treatment:	Externalizing: -14.75 (SE=8.37) Internalizing: -14.19 (SE=9.26) Global severity index: -0.03 (SE=0.497) Externalizing: 3.29 (SE=9.97) Internalizing :-6.18 (SE=9.67) Global severity index: -0.37 (SE=0.428)
Results	ICER: 1 point improvement in externalizing scores for usual care was associated with a cost of €1,561. 1 point improvement in externalizing scores for MST was associated with a costs of €404. After 12 months both treatments have comparable costs and externalizing scores.			

Studies considering costs and benefits						
Olsson[4] (2010)	**Costs intervention (MST)** Treatment costs: Travel:	 €10.789 €53 (133)	**Costs comparator (CAU)** Travel:	 €151 (225)	**Benefits intervention (MST)** Psychosocial and behavioral effects: - Social services (placement): €31,947 (€65,869) Social services (nonplacement): €8,557 (19,459) National board of institutional care (rebate): €3.009(11.014) National board of institutional care (placements): €3.593 (31.937) Wider societal costs and benefit: set to zero Psychosocial and behavioral effects: set to zero	**Benefits comparator (CAU)** Program effects: Social services (Placement): €36,707 (73,407) Social services (nonplacement): €14,914 (15,405) National board of institutional care (rebate): €2,375 (9,949) National board of institutional care (placements): 0 (0) SEK Wider societal costs and benefit: set to zero
	Results	The net loss to society after two years is €4,555				

Table 4 | Continued

Studies considering costs and benefits							
Cary (2013)	**Costs interventions (MST+YOT)**		**Costs comparator (YOT)**		**Benefits interventions (MST+YOT)**	**Benefits comparator (YOT)**	
	Treatment costs:	€3.013 (1.940)	Social worker:	€1.023 (779)	Offending behavior (Young offender information system):		
	Social worker:	€733 (446)	Reparation worker:	€83 (14)	€12,397 (18 472)	€15,409 (24,013)	
	Reparation worker:	€100 (131)	Drugs worker:	€78 (152)			
	Drugs worker:	€54 (74)	Connexions worker:	€18 (61)			
	Connexions worker:	€33 (69)	Parenting worker:	€90 (182)			
	Parenting worker	€36 (137)	Group worker:	€22 (44)			
	Group worker:	€17 (34)	Psychologist:	€30 (91)			
	Psychologist:	€17 (67)	Other appointments:	€26 (95)			
	Other appointments:	€20 (59)					
	Results	Difference (Costs+benefits) between treatments €1,612 (95% C.I-€7,699-€to 10,924) In the cost-effectiveness plane, we see, there is 63% probability that the net benefit of MST+Yot is positive in favor of the MST+YOT group.					
Dopp (2014)	**Costs interventions (MST)**		**Costs comparator (CAU)**		**Benefits intervention (MST)**		**Benefits comparator (IT)**	
	Costs per patient:	€9,756	Costs per patient:	€1,843	Benefits for taxpayer:		Benefits for taxpayer:	
					Murder:	€0	Murder:	€0
					Sexual offenses:	€922	Sexual offenses:	€602
					Robbery:	€188	Robbery:	€308
					Assault:	€1,156	Assault:	€1,697
					Property:	€2,395	Property:	€1,899
					Drug:	€916	Drug:	€1,334
					Theft:	€131	Theft:	€188
					Stolen property:	€24	Stolen property:	€53
					Fraud:	€259	Fraud:	€224
					Assault:	€236	Assault:	€294
					Drug:	€777	Drug:	€598
					TOTAL:	€7,007	TOTAL:	€7,197

Table 4 | Continued

Results	Crime victim avoided expenses		Net present values and benefit-cost ratios		Sensitivity analysis	
	Murder/manslaughter:		Net present values		Max (plausible) values	
	Tangible:	€6,125	Referred youths:		Crime victim intangible benefits:	€48,087
	Intangible:	€11,365	Taxpayer:	€2,348	Sibling juvenile arrest rates:	€30.74
	Sexual:		Crime victim tangible:	€2,389	Discount rates:	€24.063
	Tangible:	€259	Crime victim intangible	€9,375	Min (plausible) values	
	Intangible:	€3,439	Cumulative:	€29,939	Crime victim intangible benefits:	€17.561
	Robbery:		Siblings:		Sibling juvenile arrest rates:	-
	Tangible:	€575	Taxpayer:	€674	Discount rates:	€36.704
	Intangible:	€1,422	Crime victim tangible:	€2,702		
	Assault:		Crime victim intangible:	€4,533		
	Tangible:	€539	Cumulative:	€6,561		
	Intangible:	€2,926	Sibling pairs:			
	Property:		Taxpayer:	€1,399		
	Tangible:	€3,914	Crime victim tangible:	€3,499		
	Intangible:	€0	Crime victim intangible:	€11,238		
	Drug:		Cumulative*:	€31,962		
	Tangible:	€0	*: Includes the incremental costs of MST over IT			
	Intangible:	€0				
	TOTAL					
	Tangible:	€11,412				
	Intangible:	€19,151				

Table 4 | Continued

		Benefit cost ratio Referred youths Taxpayer: 1.3 Crime victim tangible: 1.3 Crime victim intangible 2.19 Cumulative: 4.78 Siblings: Taxpayer: - Crime victim tangible: - Crime victim intangible: - Cumulative: - Sibling pairs Taxpayer: 1.18 Crime victim tangible: 1.44 Crime victim intangible: 2.42 Cumulative*: 5.04 *: Includes the incremental costs of MST over CAU		
Borduin (2015)	Costs interventions (MST-PSB) Costs per patient: €10,566	Costs comparator (CAU) Costs per patient: €4,610	**Benefits intervention (MST-PSB)** Benefits for taxpayer Murder: €0 Sexual offenses: €6,419 Robbery: €2,189 Assault: €0 Property: €2,831 Drug: €1,899 Theft: €180 Stolen property €0 Fraud: €91 Assault: €250 Drug: €512 TOTAL: €14,371	**Benefits comparator (CAU)** Benefits for taxpayer Murder: €0 Sexual offenses: €15,756 Robbery: €0 Assault: €2,194 Property: €3,790 Drug: €518 Theft: €65 Stolen property: €39 Fraud: €75 Assault: €289 Drug: €112 TOTAL: €22,839

Table 4 | Continued

Results	Crime victim avoided expenses	Net present values and benefit-cost ratios	Sensitivity analysis
	Murder/manslaughter: Tangible: €41,048 Intangible: €76,169 Sexual: Tangible: €1,739 Intangible: €23,044 Robbery: Tangible: €3,850 Intangible: €9,529 Assault: Tangible: €3,612 Intangible: €19,611 Property: Tangible: €26,244 Intangible: €0 TOTAL Tangible: €76,494 Intangible: €128,353	<u>Net present values</u> Referred youths: Taxpayer: € 79,891 Crime victim tangible: €70,538 Crime victim intangible €122,397 Cumulative*: €284,739 Siblings: <u>Benefit cost ratio</u> Referred youths Taxpayer: 14.41 Crime victim tangible: 12.84 Crime victim intangible 21.55 Cumulative: 48.81 *: Includes the incremental costs of MST over CAU	Max (plausible) values: Crime victim intangible benefits: €387,085 Discount rates: €239,009 Posttreatment arrest rates: €478,277 Min (plausible) values: Crime victim intangible benefits: €188,217 Discount rates: €311,107 Posttreatment arrest rates: €91,673

Currency and price year: Schoenwald (1996).USD, 1996; French (2003). USD, 1999; Mc Collister (2009). USD,2008; Olsson(2010) SEK, 2007; Cary (2013). Pounds, 2008; Dopp (2014) USD, 2012; Borduin(2015) USD,2013.. When a price year was not stated it was estimated by taking the mean year of the study duration or when not available subtracting 1 from the year of publication of the study. For Schoenwald et al. (2006), 1996 was taken as prices year although the study was published in 1996. This was because they already published their first study in 1996 (preliminary findings) and subsequently probablythe current study was conducted in 1996.

MST=multisystemic therapy; Joint=Combination of individual and family therapy; group=skill-focused psycho-education group intervention; CAU=Care As Usual; FSN= Family support network; MDFT= multidimensional family treatment; MET/CBT12: Motivational enhancement treatment/ cognitive behavior therapy, 12 sessions; MET/CBT5= Motivational enhancement treatment/ cognitive behavior therapy, 5 sessions; Acra= Adolescent community reinforcement approach; DC=Drug Court with community services; DC+MST=Drug court with Multisystemic therapy; DC+MST+CM=Drug court with MST and enhanced with a contingency management programs; FFT= functional family therapy; FC= Family court with community service; MST-PSB MST for sexual behaviors; ICER incremental cost-effectiveness ratio.

Substance abuse

Six studies were identified which included adolescents that were treated for substance abuse [22, 24, 25, 27, 29, 32]. Three of these studies considered costs and effects [25, 27, 29] and three considered both costs and benefits [22, 24, 32].

Studies considering costs and effects

In the study of French et al. [25] FFT was shown to be more cost-effective than a skill-focused psycho-education group intervention for treating substance use disorders and delinquency after the first 4 months. After 12 months no such effect was observed. Therefore, after 12 months the cost-effectiveness analysis reduced to a simple cost minimization analysis. As only treatment costs were considered (narrow perspective), the intervention with the lowest intervention costs, in this case group therapy, was considered to be economically beneficial. In another study, Dennis et al. [29] computed cost-effectiveness ratios and these ratios indicated that overall, the most cost-effective interventions were Motivational Enhancement Treatment/ Cognitive Behavior Therapy, 5 sessions (MET/CBT5) and Motivational enhancement treatment/ Cognitive Behavior Therapy, 12 sessions (MET/CBT12) when compared to Family Support Network (FSN) in Trial 1 and Adolescent Community Reinforcement Approach (ACRA) and MET/CBT5 when compared to MDFT in Trial 2. Sheidow et al. [27], computed Average Cost-Effectiveness Ratios (ACERS). ACERS only incorporate the pre-post treatment effect of one single treatment so treatments are not directly compared. Although this study showed that Drug Court with community services (DC) was more cost effective compared to FC regarding substance use disorders and that the addition of multi-systemic therapy (MST) resulted in an economically more beneficial treatment, the treatments were not directly compared [27].

Studies considering costs and benefits

Three of the studies that considered adolescents with substance use disorders, considered costs and benefits [22, 24, 32]. The study of French et al. [22] indicated that MET/CBT-5, MET/CBT-12 and FSN generated significant economic benefits to society for substance abusing adolescents, MDFT and ACRA did not generate these benefits. McCollister et al. [24] showed that the savings in costs offset the treatment costs of DC, especially for DC/MST/ CM, in juvenile drug court participants when compared to FC (Family court with community services). Schoenwald [32] showed that the monetary benefits of MST compared to CAU for substance use disorder almost offset the higher costs of MDFT. Over time the difference between benefits and costs may be reduced to a complete offset.

Delinquency/externalizing disorders

Five studies considered adolescents with delinquency or externalizing disorders; the study of Sheidow et al. [23], Olsson [26], Cary et al. [28], Dopp et al. [30] and Borduin et al. [31] respectively included patients with a psychiatric crisis, patients with a conduct disorder, delinquent adolescents, serious juvenile offenders and juvenile sexual offenders. One study, Sheidow et al. [23], considered both costs and effects and four studies [26, 28, 30, 31] considered both costs and benefits.

Studies considering costs and effects

In the study of Sheidow et al. [23], MST was effective in the short term (4 months) in terms of externalizing behavior compared to care as usual for patients with psychiatric emergencies. But MST appeared equally effective on the cost measure over the long term (12 months).

Studies considering costs and benefits

Olsson [26] showed that for adolescents with conduct disorder MST's benefits did not offset the costs and that MST was subsequently associated with a net loss to society. The study of Cary et al. [28] showed that MST in combination with CAU has a scope to generate cost savings when compared to providing CAU alone. The cost-benefit study of Dopp et al. [30] indicated that MST, when delivered to serious juvenile offenders, produces economic benefits well into adulthood. Borduin et al. [31] showed that when juvenile sexual offenders are treated with MST-PSB; this treatment can produce lasting economic benefits.

Quality of the studies

Only for one study [23] commentary was available from the NHS-EED. We compared the commentary on the study with our quality assessment checklists to evaluate if all issues were addressed. The quality of the studies was not only assessed for the 7 unique studies but for the 9 studies. The argument for including all studies was to differentiate between methods (e.g. analysis), display of results and discussion even though they were based on the same study. The quality assessed with the BMJ checklist was between 52 and 86% (Table 3.5). The quality assessed with the CHEC list was between 50 and 79% (Table 3.5). Up to date, there are no thresholds (minimum number of criteria satisfied) for these checklists to determine the difference between bad and good quality economic evaluations [18]. Overall, the outcomes on the checklists matched although quality assessed with the CHEC list was consequently lower. The largest difference in quality

percentages was 20%. All studies clearly stated their primary outcome measures. Most studies did not report all relevant costs and effects.

Table 5 | Assessments of the quality of the studies with the Drummond checklist and the CHEC list

British Medical Journal Checklist
1. The research question is stated.
2. The economic importance of the research question is stated.
3. The viewpoint(s) of the analysis are clearly stated and justified.
4. The rationale for choosing alternative programmes or interventions compared is stated.
5. The alternatives being compared are clearly described
6. The form of economic evaluation used is stated.
7. The choice of form of economic evaluation is justified in relation to the questions addressed.
8. The source(s) of effectiveness estimates used are stated.
9. Details of the design and results of effectiveness study are given (if based on a single study).
10. Details of the methods of synthesis or meta-analysis of estimates are given (if based on a synthesis of a number of effectiveness studies).
11. The primary outcome measure(s) for the economic evaluation are clearly stated.
12. Methods to value benefits are stated.
13. Details of the subjects from whom valuations were obtained were given.
14. Productivity changes (if included) are reported separately.
15. The relevance of productivity changes to the study question is discussed.
16. Quantities of resource use are reported separately from their unit costs.
17. Methods for the estimation of quantities and unit costs are described.
18. Currency and price data are recorded.
19. Details of currency of price adjustments for inflation or currency conversion are given.
20. Details of any model used are given
21. The choice of model used and the key parameters on which it is based are justified.
22. Time horizon of costs and benefits is stated.
23. The discount rate(s) is stated.
24. The choice of discount rate(s) is justified.
25. An explanation is given if costs and benefits are not discounted.
26. Details of statistical tests and confidence intervals are given for stochastic data.
27. The approach to sensitivity analysis is given.
28. The choice of variables for sensitivity analysis is justified.
29. The ranges over which the variables are varied are justified.
30. Relevant alternatives are compared.
31. Incremental analysis is reported.
32. Major outcomes are presented in a disaggregated as well as aggregated form
33. The answer to the study question is given.
34. Conclusions follow from the data reported.
35. Conclusions are accompanied by the appropriate caveats.
Total score British medical journal checklist (%)

1*	2*	3*	4*	5*	6*	7*	8*	9*	10*	11*
-	-	✔	✔	✔	✔	✔	-	✔	✔	✔
✔	-	✔	✔	-	✔	✔	✔	✔	✔	✔
-	✔	-	✔	-	-	✔	✔	-	-	-
✔	-	-	-	-	-	✔	-	✔	-	-
✔	✔	✔	✔	-	-	-	-	✔	✔	✔
-	✔	✔	✔	-	✔	✔	✔	✔	✔	✔
NC	✔	✔	✔	-	✔	✔	✔	✔	✔	✔
✔	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔
✔	NA	✔	✔	✔	✔	✔	✔	-	✔	✔
NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
✔	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔
✔	✔	NA	✔	✔	NA	✔	NA	✔	✔	✔
✔	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔
NA	✔	NA	NA	NA	NA	NA	NA	NA	-	-
-	-	-	-	-	-	✔	-	-	-	-
✔	✔	-	-	-	-	-	-	✔	✔	✔
-	-	-	-	✔	-	✔	✔	✔	✔	✔
✔	✔	-	✔	-	-	-	✔	✔	✔	✔
✔	✔	-	-	-	-	✔	-	✔	✔	✔
NA	✔	✔	✔	✔	✔	NA	NA	✔	NA	NA
NA	-	-	✔	-	-	NA	NA	-	NA	NA
✔	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔
NA	NA	NA	NA	NA	NA	✔	NA	✔	✔	✔
NA	NA	NA	NA	NA	NA	✔	NA	✔	✔	✔
NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
-	-	✔	-	✔	-	✔	✔	✔	-	-
✔	-	✔	-	-	-	✔	NC	✔	✔	✔
✔	NA	NA	NA	NA	NA	✔	NA	✔	✔	✔
NC	NA	NA	NA	NA	NA	✔	NA	✔	✔	✔
✔	NC	-	NC	✔	NS	✔	✔	✔	✔	✔
✔	✔	-	✔	✔	-	✔	-	✔	✔	✔
✔	✔	✔	-	✔	✔	✔	✔	✔	✔	✔
✔	NC	✔	✔	✔	✔	✔	✔	✔	✔	✔
✔	✔	✔	✔	✔	✔	✔	✔	✔	-	-
-	-	✔	✔	-	✔	✔	✔	-	✔	-
68	61	63	68	54	52	86	70	83	81	77

Table 5 | Assessments of the quality of the studies with the Drummond checklist and the CHEC list. Continued.

CHEC list
1. Is the study population clearly described?
2. Are competing alternatives clearly described?
3. Is a well-defined research question posed in answerable form?
4. Is the economic study design appropriate to the stated objective?
5. Is the chosen time horizon appropriate to include relevant costs and consequences?
6. Is the actual perspective chosen appropriate?
7. Are all important and relevant costs for each alternative identified?
8. Are all costs measured appropriately in physical units?
9. Are costs valued appropriately?
10. Are all important and relevant outcomes for each alternative identified?
11. Are all outcomes measured appropriately?
12. Are outcomes valued appropriately?
13. Is an incremental analysis of costs and outcomes of alternatives performed?
14. Are all future costs and outcomes discounted appropriately?
15. Are all important variables, whose values are uncertain, appropriately subjected to sensitivity analysis?
16. Do the conclusions follow from the data reported?
17. Does the study discuss the generalizability of the results to other settings and patient/client groups?
18.Does the article indicate that there is no potential conflict of interest of study researcher(s) and funder(s)?
19.Are ethical and distributional issues discussed appropriately?
Total score CHEC** (%)

*Studies: Schoenwald et al., 1996; 2 French et al., 2003; 3 Sheidow et al., 2004; 4 Dennis et al., 2004); 5 McCollister et al., 2009; 6 French et al., 2008; 7 Olsson, 2010; 8 Sheidow et al., 2012; 9 Cary et al., 2013; 10 Dopp et al., 2014; 11. Borduin et al., 2015. NS: Not stated; NA: Not applicable; NC: Not clear. Explanation criteria checklist: British medical journal checklist: 1.A specific question is not necessary, as long as the goal of the research is clearly stated; 5. The competing alternatives may also be described in a different accessible paper from the RCT in more detail 10. The presentation of the results is clearly given and discussions of the study contain generalizability and comparison with other studies. CHEC list: 5: Chosen time horizon is appropriate when after a certain time no additional effects are attained. **Scores were calculated by dividing the positively checked items on the quality checklist by the total minus items on the checklist that were not applicable (NA) to the study

	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔
	✔	✔	✔	✔	-	-	-	-	✔	✔	✔
	-	-	✔	✔	✔	✔	✔	-	✔	✔	✔
	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔
	NS	NS	✔	NS	NS	NS	✔	NS	NS	✔	✔
	-	✔	-	-	-	-	✔	-	-	-	-
	-	-	NS	-	-	-	-	-	-	✔	✔
	✔	✔	-	-	-	-	✔	-	✔	✔	✔
	✔	✔	-	✔	✔	NS	✔	✔	✔	✔	✔
	-	-	✔	✔	-	✔	✔	✔	✔	✔	✔
	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔
	-	✔	✔	✔	✔	✔	-	✔	✔	✔	✔
	✔	✔	-	✔	✔	-	✔	-	✔	✔	✔
	NA	NA	NA	NA	NA	NA	✔	NA	✔	✔	✔
	✔	-	-	-	-	-	✔	-	✔	✔	✔
	✔	✔	✔	✔	✔	✔	✔	✔	✔	-	-
	-	-	✔	-	-	✔	✔	✔	-	-	-
	-	✔	-	-	-	-	-	-	-	-	-
	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔
	56	67	61	61	50	50	79	50	79	74	74

Risk of bias

All studies were RCTs [22-32]. Two of these studies [23, 32] only included patients receiving Medicaid (an aid program regarding insurances for low income families in the United States). For these studies, the RCT of the effect study contained (due to randomization) balanced samples. However, these samples were not checked for balance after the selection of participants who received medicaid, so they were at risk for selection bias. All studies had a high risk of performance bias, as blinding of both therapist and patient is impossible. For two studies [23, 32] blinding was not necessary as both the cost and outcome data were extracted from existing data systems (The medicaid billing records). Although blinding of outcome assessors is possible to reduce detection bias, no study reported to have done so. Blinding is also necessary for pre-allocation assessment. All studies were based on randomized controlled trials where allocation concealment is necessary. The studies included in this review, did not explicitly refer to the allocation concealment. Three studies were at risk of attrition bias. These three studies did not describe the number of patients that dropped out from the study [24, 27, 32]. Two studies only described the overall attrition rate [22, 25]. For one study [29] however, overall attrition rate could be extracted by using the study of French et al. [22] as it was based on the same participants. Dropout in the effect-study of Sheidow et al. [23] was low and although no dropout was described for the economic evaluation, as the economic evaluation is based on the same participants, this is expected to be low. Overall, dropout rate (when measured) seemed low. Reporting bias was assessed by reading protocols from the studies and no bias was reported. Only for two studies [22, 29] a protocol existed. Other studies did not have such a protocol, although for three studies trial registrations were present [24, 27, 28]. There were no indications of deviations from the original design. The economic evaluations did not always include all clinical outcomes that were available [23-26, 32] as there was often only interest in specific outcomes. One study [25] excluded clinical outcomes as there was no difference between treatments in terms of outcomes and so only costs were considered (costs minimization). The exclusion of outcomes was not related to possible negative impact on the results as effects in the studies were equally or more beneficial when compared to the effects of the comparator.

Methodological summary

Uncertainty around treatment costs was not presented in four studies as averages of these costs were used [24, 27, 30, 31]. In six studies [22, 23, 25, 29-32]

uncertainty around the (other) estimates was not (fully) addressed. In seven studies, a simple one-way sensitivity analysis was used to assess the impact that changes in a certain parameter will have on the conclusions [22, 23, 26, 28, 30-32]. In two studies, sensitivity analysis was applied by imputing missing data in different ways. Outcomes proved to be robust [27, 28]. Two studies performed scenario analyses meaning that cost estimates (surrounded by uncertainty) were increased or decreased. Data proved to be robust [26, 32]. In another study a sensitivity analysis was carried out to assess the effect, which outliers in each therapy group had on outcomes, but this did not have an effect the results. In the studies of Dopp et al. [31] and Borduin et al. [31] a sensitivity analysis was applied by using plausible minimum and maximum values (obtained from other studies) for offense categories, arrest rates and discount rates. French et al. [22] used different models, which assessed the effect on using more or less covariates in the models but it did not affect the results. In six of the studies cost-effectiveness/utility/benefits were assessed based on models [22-25, 28, 32]. Four of these studies used simple regression models [23-25, 28] and two used a more advanced least squares random effect model [19, 26]. The remaining three studies did not integrate any model in the analysis. Three studies did not report their price year (the year to which costs are indexed) [23, 24, 32]. Authors of three studies indicated that a societal perspective was adopted, where not only health care costs but also other costs, for example those associated with lost or impaired ability to work, were taken into account [22, 26, 29]. However, this was only true for the study of Olsson [26], as this was the only study to assess costs outside the health care sector. In the studies of Dennis et al. [29] and French et al. [22], the societal part was defined as using market values for calculating the costs of goods and services used. Dopp et al. [30] and Borduin et al. [31] conducted cost-benefit analyses and did not explicitly mention their perspective. Both studies focused on taxpayer benefits and expressed intangible benefits in monetary values. Cary et al. [28] used a narrow perspective as only services that were recorded by a specific data-system were included (appointments with social workers, connexion workers (a United Kingdom (UK) governmental information, advice, guidance and support service for young people aged thirteen to nineteen), reparation workers (coordinates and supports a range of interventions and community reparation projects that young people will have to undertake as part of their Referral or Community Order), parenting workers, group workers and psychologists). Sheidow [23] adopted the perspective of an institution. Other studies did not explicitly state their perspective. Most of the studies only reported treatment costs. A summary of the costs and clinical

outcomes measured in the studies is provided in Table 3.6. Following Drummond et al. [33], full economic evaluations should not only report costs, but also health outcomes. Four studies were classified as cost-effectiveness analyses [23, 25, 27, 29]. Only one of these studies compared treatments using an incremental cost-effectiveness ratio [29] as described for instance by Drummond et al. [33]. The cost-effectiveness analysis of French et al. [25] was reduced to a simple cost minimization analysis as the effects of both treatments after analysis proved to be similar. Sheidow et al. [27] calculated average cost-effectiveness ratios (ACER), which means that there was no direct comparison between treatments but only between the before- and after treatment costs and effects of every participant. In four studies it was explicitly stated that cost-benefit analyses [22, 26, 30,31] were performed. Olsson [26] considered psychosocial and behavioral effects, but as no difference was observed regarding these clinical measures between treatments, these effects were excluded from the analysis. French et al. [22] did not value the health outcomes on which the intervention was focused (like reduction in days of substance use) but did value the effects of treatment on education, employment and criminal activity. Dopp et al. [30] and Borduin et al. [31] conducted a cost-benefit analysis; the cost outcome were the treatment costs and the benefits were defined as taypayer benefits, tangible benefits and intangible benefits were expressed in monetary values. Cary et al. [28] classified his study as a cost-offset evaluation. He calculated the net-benefit, but stated that his study cannot be viewed as a cost-effectiveness study as he did not measure health outcome. Two studies did not state the type of economic analyses they performed [24, 32], but did consider both costs and benefits. Mcollister [24] indicated that her study was not a full economic evaluation, as she only considered treatment costs. This is also the case concerning the study of Sheidow et al. [27], however, this study was stated to be a cost-effectiveness analysis. Furthermore, Schoenwald et al. [32] did not classify their study explicitly but considered both costs of different health care services and monetary benefits so it can be considered a cost-benefit analysis.

Table 6 | Overview of costs and clinical outcome measures used in studies

	Treatment costs	Other health-care costs	Costs outside health care sector	Perspective used in the economic evaluations	Clinical outcome measure
(Schoenwald et al., 1996)	✔	✔		Healthcare	-
(French et al., 2003)	✔			Institution	-
(Sheidow et al., 2004)	✔	✔		Healthcare	CBCL/GSI
(Dennis et al., 2004)	✔			Institution	-
(McCollister et al., 2009)	✔			Institution	SRD
(French et al., 2008)	✔			Institution	YSR/days of marijuana use
(Olsson, 2010)	✔		✔	Societal	-
(Sheidow et al., 2012)	✔			Institution	TLFB/SRD
(Cary et al., 2013)	✔			Institution	-
(Dopp et al. 2014)	✔		✔	Societal	-
(Borduin et al. 2015)	✔		✔	Societal	-

CBCL=Child Behavior Checklist; GSI: Global severity index; SRD=Self-Report Delinquency Scale; TLFB=Timeline Follow-back Form; YSR=Youth Self Report

Limitation/generalizability summary

Four studies commented on their generalizability [23, 25-27]. Sheidow et al. [23] reported that as their sample only consisted of youths enrolled in Medicaid, which are generally economically less advantaged, findings cannot be generalized to a more economically advantaged population. The same is true, although not stated, for the study of Schoenwald et al. [32] who also analyzed Medicaid data. The study of Olsson [26] was conducted in Sweden, where MST is twice more expensive than in the USA and may play a different role in society. MST in Sweden may be used as an alternative to nonplacement interventions as opposed to an alternative to placement interventions as found in other studies. Also in the study of French et al. [25], which was conducted in Mexico, location and small sample size were indicated as limitations for generalizability. The same was true, although not stated, for the study of Cary et al. [28] which was conducted in the United Kingdom. Also an important limitation (but not mentioned as such) were

the omissions of uncertainty around the estimates in the studies of Dopp et al. [30] and Borduin et al. [31], so the results should be interpreted with caution. Furthermore, the study of Borduin et al. [31] was based on a very small (the smallest one in this review) sample size (only 48 patients) so uncertainty around the estimates (not reported) is expected to be high. Sensitivity analysis is not a solution for this problem as significance of the results cannot be determined (as the estimates in the sensitivity analysis are also subjected to uncertainty). The juvenile drug court programs, analyzed in the study of Sheidow et al. [27] are not easily generalized to other settings as they show great variation due to absence of a strict format. In addition, other settings may have different populations and salaries implying differences in costs. Almost all studies were cautious with drawing conclusions on their data. They not only recognized limitations within their research but also recognized that the number of economic evaluations is very limited and more research is needed before being able to draw conclusions [22-28, 32].

Meta analysis

The data from the economic evaluations were not pooled as the population, setting, outcomes, costs and interventions were not comparable across studies.

Discussion

This systematic review summarized and evaluated the cost-effectiveness of family/family-based therapy for adolescents with externalizing disorders, substance use disorder and delinquency. The overall quality of these studies was low; they produced mixed results. Research should consider a wider perspective and take into account all relevant costs and effects using sophisticated models. Studies evaluating family/family-based therapy concerned various outcomes and costs, and investigated a variety of treatments in various populations in different settings. Therefore it was not possible to conduct a meta-analysis. As expected, most of the studies were conducted in the United States where family/family-based treatments originate from [10, 11, 34]. The findings cannot be easily generalized to other health care systems as they differ between countries. The quality assessments showed that overall studies scored between 50 and 86% and only two studies scored higher than 80% [26, 28, 30, 31]. Studies that were conducted more recently, were in general higher of quality. When the two most recent studies [30, 31] were not considered, the quality of the studies overall was slightly higher for those studies originating from Europe. The quality of the two most recent studies was high when using the quality checklists, however, they

also contained some important limitations. Firstly, although quality checklists only contain one question with respect to uncertainty around the estimates, it can be of paramount importance, especially when the sample size is low. Secondly, these studies are not easily generalized to a European setting as they conducted cost-benefit analyses, opposed to cost-effectiveness analyses that are commonly applied in European studies. Although the checklists used to assess quality of the studies depend on the subjective evaluation of the researchers and have yet not been validated, these two checklists have received much scrutiny and are therefore recommended [18]. Recommendations that follow from the quality assessment of the studies that were included in the review, are the following. Different treatments that are included in the study should be described more clearly so the differences and similarities between treatments are understandable. In many of the studies included in the review, the perspective taken was not mentioned or did not match with the categories of the costs that were included. In line with guidelines for economic evaluations the perspective should be stated [33]. A more broad perspective (societal versus healthcare) is recommended. The unit costs and resource use should be reported separately and a source of the references for the unit costs should be given. It is also important to explicitly mention whether a study is considered a cost-effectiveness/cost-benefit or cost-utility analysis. Most studies included in the review used no model or simple models (regression). More complex models, like multilevel analysis, should be used. In this way covariates can be included, correlation between measurements over time can be addressed, missing data is accounted for and skewness in the costs and effects is considered. Uncertainty around costs should also be presented by using for instance bootstrapped costs/effects confidence intervals and can be visualized in a cost-effectiveness plane. Sensitivity analysis should be applied to variables that are uncertain (the rationale behind it should be explained). A one way sensitivity analysis is not always sufficient and a sensitivity analysis also taking into account interactions between variables should be considered. A common discount rate should be applied for all costs and effects. Summary measures of the cost-benefit, cost-effectiveness or cost utility should be given. In case of a cost effectiveness analysis incremental cost-effectiveness ratio (ICERS) should be calculated. For conducting economic evaluations it is advised to consult a health economist.

Conclusions

Although family/family-based treatments are widely used and can be considered as effective for the treatment of a wide range of disorders [17], cost-effectiveness also needs to be addressed. Taking cost-effectiveness into account may have a large impact as family/family-based treatments are expensive. This review has summarized the economic evidence of family/family-based therapy for substance use disorders and delinquency in adolescents in a systematic and transparent way by using state of the art guidelines [18, 19]. As there are few studies evaluating the cost-effectiveness of family/family-based therapy and the quality of the existing studies is limited, new studies using higher quality standards are necessary. Large-scale implementation of these treatment models should be held back, until more evidence is available.

References

[1] von Sydow K, Retzlaff R, Beher S, Haun MW, Schweitzer J. The efficacy of systemic therapy for childhood and adolescent externalizing disorders: A systematic review of 47 RCT. Family Process. 2013;52(4):576-618.

[2] Baldwin SA, Christian S, Berkeljon A, Shadish WR. The effects of family therapies for adolescent delinquency and substance abuse: A meta-analysis. J Marital Fam Ther. 2012;38(1):281-304.

[3] Broman CL, Reckase MD, Freedman-Doan CR. The role of parenting in drug use among black, latino and white adolescents. J Ethn Subst Abuse. 2006;5(1):39-50.

[4] Choquet M, Hassler C, Morin D, Falissard B, Chau N. Perceived parenting styles and tobacco, alcohol and cannabis use among french adolescents: Gender and family structure differentials. Alcohol Alcohol. 2008;43(1):73-80.

[5] Hoeve M, Dubas JS, Gerris JR, van der Laan PH, Smeenk W. Maternal and paternal parenting styles: Unique and combined links to adolescent and early adult delinquency. J Adolesc. 2011;34(5):813-827.

[6] Kristjansson AL, Sigfusdottir ID, Allegrante JP. Adolescent substance use and peer use: A multilevel analysis of cross-sectional population data. Subst Abuse Treat Prev Policy. 2013;8:27-597X-8-27.

[7] Curcio AL, Mak AS, George AM. Do adolescent delinquency and problem drinking share psychosocial risk factors? A literature review. Addict Behav. 2013;38(4):2003-2013.

[8] Liddle HA, Rowe CL, Dakof GA, Ungaro RA, Henderson CE. Early intervention for adolescent substance abuse: Pretreatment to posttreatment outcomes of a randomized clinical trial comparing multidimensional family therapy and peer group treatment. J Psychoactive Drugs. 2004;36(1):49-63.

[9] Carr A. The effectiveness of family therapy and systemic interventions for child-focused problems. Journal of Family Therapy. 2008;31(1):3-4.

[10] Henggeler SW, Borduin CM. Family therapy and beyond: A multisystemic approach to treating the behaviour problems of children and adolescents. Pacific Grove: CA: Brooks/Cole; 1990.

[11] Alexander JF, Parsons BV. Functional family therapy. Monthery: CA: Brooks/Cole; 1982.

[12] Liddle HA. Multidimensional Family Therapy for Adolescent Cannabis users, Cannabis Youth Treatment (CYT) Series, 2002;5.

[13] Retzlaff R, von Sydow K, Beher S, Haun MW, Schweitzer J. The efficacy of systemic therapy for internalizing and other disorders of childhood and adolescence: A systematic review of 38 randomized trials. Fam Process. 2013;52(4):619-652.

[14] Hendriks V, van der Schee E, Blanken P. Treatment of adolescents with a cannabis use disorder: Main findings of a randomized controlled trial comparing multidimensional family therapy and cognitive behavioral therapy in the netherlands. Drug Alcohol Depend. 2011;119(1-2):64-71.

[15] Van der Stouwe T, Asscher JJ, Stams GJ, Deković M, van der Laan PH. The effectiveness of multisystemic therapy (MST): A meta-analysis. Clinical psychology review. 2014;34(6):468-481.

[16] Darnell AJ, Schuler MS. Quasi-experimental study of functional family therapy effectiveness for juvenile justice aftercare in a racially and ethnically diverse community sample. Child Youth Serv Rev. 2015;50:75-82.

[17] Morgan TB, Crane DR. Cost-effectiveness of family-based substance abuse treatment. J Marital Fam Ther. 2010;36(4):486-498.

[18] Shemilt I, Mugford M, Byford S, et al. Incorporating economics evidence. In: Higgins J.P.T. GS, ed. Cochrane handbook for systematic reviews of interventions. 5.1.0. ed. handbook.cochrane.org.: The Cochrane Collaboration; 2011.

[19] Liberati A, Altman DG, Tetzlaff J, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: Explanation and elaboration. BMJ. 2009;339:b2700.

[20] Drummond MF, Jefferson TO. Guidelines for authors and peer reviewers of economic submissions to the BMJ. The BMJ Economic Evaluation Working Party. BMJ. 1996;313(7052):275-83.

[21] Evers S, Goossens M, de Vet H, van Tulder M, Ament A. Criteria list for assessment of methodological quality of economic evaluations: Consensus on health economic criteria. Int J Technol Assess Health Care. 2005;21(2):240-245.

[22] French MT, Roebuck MC, Dennis ML, Godley SH, Liddle HA, Tims FM. Outpatient marijuana treatment for adolescents. economic evaluation of a multisite field experiment. Eval Rev. 2003;27(4):421-459.

[23] Sheidow AJ, Bradford WD, Henggeler SW, et al. Treament costs for youths receiving multisystemic therapy or hospitalization after a psychiatric crisis. Psych Serv. 2004;55(5):548-554.

[24] McCollister KE, French MT, Sheidow AJ, Henggeler SW, Halliday-Boykins CA. Estimating the differential costs of criminal activity for juvenile drug court participants: Challenges and recommendations. J Behav Health Serv Res. 2009;36(1):111-126.

[25] French MT, Zavala SK, McCollister KE, Waldron HB, Turner CW, Ozechowski TJ. Cost-effectiveness analysis of four interventions for adolescents with a substance use disorder. J Subst Abuse Treat. 2008;34(3):272-281.

[26] Olsson TM. MST with conduct disordered youth in sweden: Costs and benefits after 2 years. Res Social Work Prac. 2010;20(2):561-571.

[27] Sheidow AJ, Jayawardhana J, Bradford WD, Henggeler SW, Shapiro SB. Money matters: Cost effectiveness of juvenile drug court with and without evidence-based treatments. J Child Adolesc Subst Abuse. 2012;21(1):69-90.

[28] Cary M, Butler S, Baruch G, Hickey N, Byford S. Economic evaluation of multisystemic therapy for young people at risk for continuing criminal activity in the UK. PLoS One. 2013;8(4):e61070.

[29] Dennis M, Godley SH, Diamond G, et al. The cannabis youth treatment (CYT) study: Main findings from two randomized trials. J Subst Abuse. 2004;27:197-213.

[30] Dopp AR, Borduin CM, Wagner DV, Sawyer AM. The economic impact of multisystemic therapy through midlife: A cost-benefit analysis with serious juvenile offenders and their siblings. J Consult Clin Psychol. 2014;82(4):694-705.

[31] Borduin CM, Dopp AR. Economic impact of multisystemic therapy with juvenile sexual offenders. J Fam Psychol. 2015;29(5):687-696.

[32] Schoenwald SK, Ward DM, Henggeler SW, Pickrel SG, Patel H. Multisystemic therapy treatment of substance abusing or dependent adolescent offenders: Costs of reducing incarceration, inpatient and residential treatment. J Child Fam Stud. 1996;5(4):431-444.

[33] Drummond MF, Sculpher MJ, Torrance GW, O'Brien BJ, Stoddart GL, eds. Methods for the economic evaluation of health care programs. 3rd ed. Oxford: Oxford University press; 2005.

[34] Liddle HA. A multidimensional model for treating the adolescent who is abusing alcohol and other drugs. In: Snyder W, Ooms T, eds. Empowering families, helping adolescents: Family-centered treatment of adolescents with alcohol, drug abuse and other mental health problems. 1st ed. Washington, DC: United States Public Health Service; 1991.Appendix 1

Appendix 1

Table 7 | Description family/family-based interventions

Family/Family based interventions	
MST	Target family interaction and the extended social systems in youths with substance abuse problems, delinquency or antisocial behavior / Permits separate meetings adolescent but preference for family /More focus on antisocial behavior/ focused both on family functioning and on extra familial functioning / Treatment team not actively involved as observers and actors but team is only self-reflexive/ Treatment team actively involved as observers and actors /degree of severity higher and combination of more problems
FFT	Target family interaction and the extended social systems in youths with substance abuse problems, delinquency or antisocial behavior/ Almost no separate meetings adolescent /More focus on antisocial behavior/More focused on family functioning less on extra familial functioning/ Treatment team not actively involved as observers and actors but team is only self-reflexive/ explicitly emphasizes therapist is integral part of the system/degree of severity lower
MDFT	Target family interaction and the extended social systems in youths with substance abuse problems, delinquency or antisocial behavior/ Separate meetings adolescent/ Focus on substance abuse / focused both on family functioning and on extra familial functioning /Treatment team not actively involved as observers and actors but team is only self-reflexive/degree of severity higher

Sources: Leukehof et al, 2008 ; Oudhof et al, 2009;
Legend: MST=multisystemic therapy; FFT= functional family therapy; MDFT= multidimensional family treatment

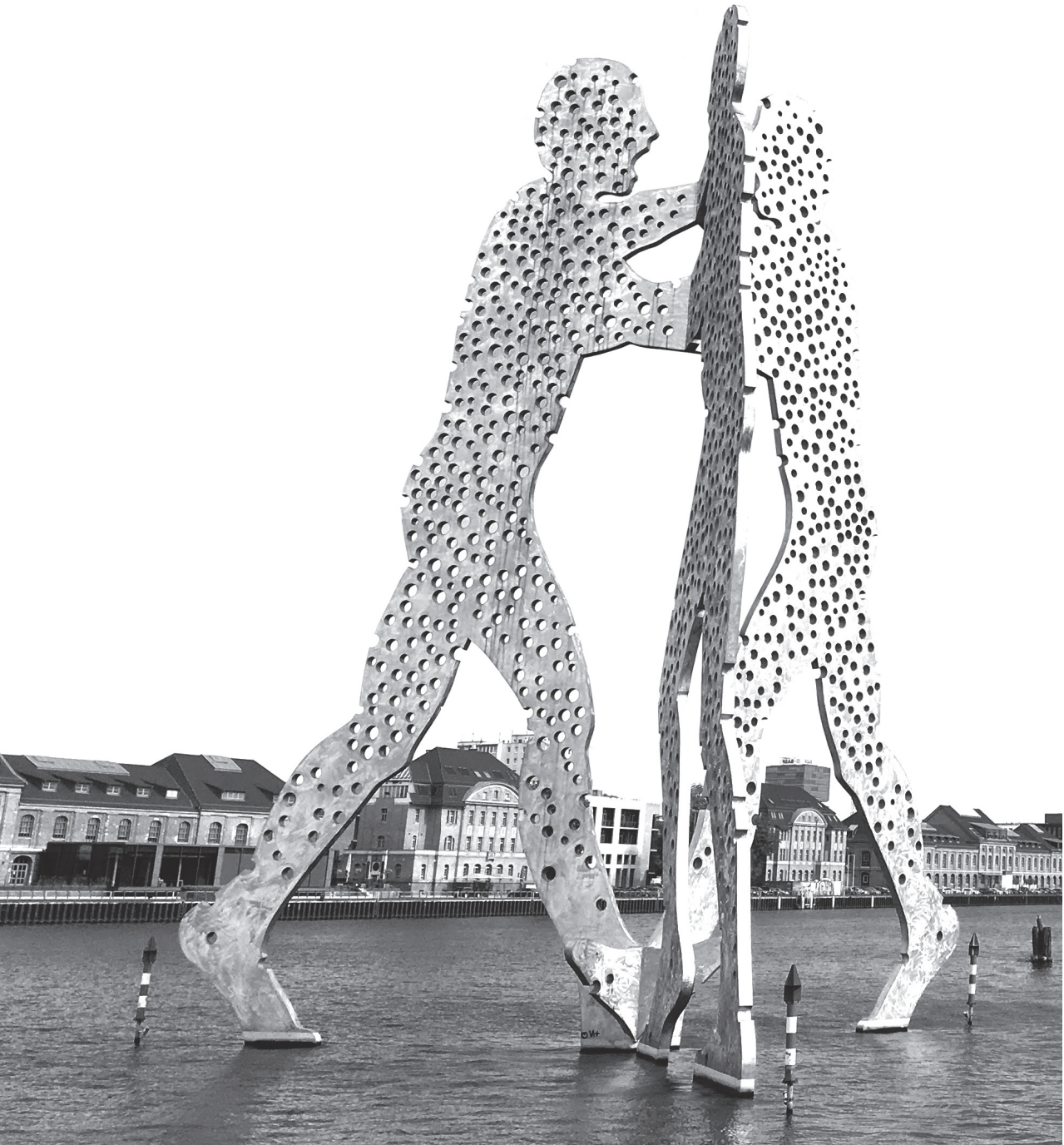
Appendix 2

Search terms Pubmed

"family therapy"[MESH]
"Functional family therapy"
(FFT NOT ("fast Fourier transform" OR "freedom-from-transfusion" OR "fast Fourier transforms" OR "fast Fourier transformation" OR "Far-Field Transform"))
"Multisystemic Therapy"
(MST NOT ("microbial source tracking" OR "minimum spanning tree"))
"Multidimensional Treatment Foster Care"
"MTFC"
"multidimensional family therapy"
"MDFT"
"family behavior therapy"
"FBT"
brief strategic family therapy"
"BSFT"
"family based therapy"[Title/Abstract]
"family based interventions"[Title/Abstract]
"family based intervention"[Title/Abstract]
"family systems intervention" [Title/Abstract]
"family systems interventions" [Title/Abstract]

"family system intervention" [Title/Abstract]
"family system interventions" [Title/Abstract]
"family intervention program"[Title/Abstract]
"family intervention programs"[Title/Abstract]
"systemic Therapy" [Title/Abstract]
OR 1-23
"economic evaluation" [title/Abstract]
"economic evaluations" [title/Abstract]
"cost effective" [title/Abstract]
"cost effectiveness" [title/Abstract]
"cost utility analysis" [title/Abstract]
"costs" [Title/Abstract] AND "effect"[Title/Abstract]
"cost" [Title/Abstract] AND "effect"[Title/Abstract]
"cost" [Title/Abstract] AND "effects"[Title/Abstract]
"costs" [Title/Abstract] AND "effects"[Title/Abstract]
"costs"[Title/Abstract] AND "benefits"[Title/Abstract]
"cost" [Title/Abstract] AND "benefit"[Title/Abstract]
"costs" [Title/Abstract] AND "benefit"[Title/Abstract]
"cost" [Title/Abstract] AND "benefits"[Title/Abstract]
"costs" [Title/Abstract] AND "utility"[Title/Abstract])
"cost" [Title/Abstract] AND "utility"[Title/Abstract])
"cost" [Title/Abstract] AND "utilities"[Title/Abstract]
"costs" [Title/Abstract] AND "utilities"[Title/Abstract])
"Cost Analysis" [title/Abstract]
"Cost Measures" [title/Abstract]
"cost benefit analysis"[title/Abstract]
"cost measure" [title/Abstract]
"cost" [title]
"costs" [title]
"cost benefit analysis" [MESH]
OR 25-48
NOT (cancer[Title/Abstract]OR psoriasis[Title/Abstract]OR "radiation therapy"[Title/Abstract] OR diabetes[Title/Abstract] OR diabetic[Title/Abstract] OR obesity [Title/Abstract] OR aids[Title/Abstract] OR HIV[Title/Abstract] OR sarcomas[Title/Abstract] OR chemotherapy[title/Abstract]))
24 AND 49 AND 50
Search terms Eric, Psycinfo and Cochrane

In Eric, the same search terms were used except for the MESH terms. In psycinfo, the MESH terms were replaced with APA's thesaurus of Psychological index Terms and in cochrane, the same terms were used.



Chapter 4

Framework for Modeling the Cost-effectiveness of Systemic Interventions Aimed to Reduce Youth Delinquency

Based on Schawo, S., van Eeren, H., Soeteman, D., van der Veldt, M.-C., Noom, M., Brouwer, W., Van Busschbach, J., Hakkaart-van Roijen, L.

The Journal of Mental Health Policy and Economics. 2012;15: 187-196.

Introduction

Child delinquency poses a high economic burden on society [1] Therefore, crime prevention and treatment of youth delinquents is of great importance to governments, in particular for Justice Departments. Systemic interventions, for instance Multisystemic Therapy (MST), Functional Family Therapy (FFT) or Parent Management Training Oregon (PTMO), are relatively costly interventions in youth health care aiming to reduce delinquent behavior [2]. Cost-effectiveness studies are still limited in the field of youth health care. However, these costly systemic family interventions compete with medical treatments and other interventions for health care budgets, increasing the need for knowledge regarding the operationalization of economic evaluations in this context.

In the Netherlands, as part of an ongoing nationwide action plan of the Ministry of Justice, recently a selection was made of evidence-based treatments for delinquent youth [3], among which MST, FFT and PMTO were implemented given their apparent effectiveness in reducing criminal activity in youths. The aim of these systemic interventions is not primarily to produce health in the sense of physical health and absence of disease, as measured in the Quality Adjusted Life Years (QALY) outcome. These interventions attempt to improve family functioning and may even intervene with the peers and school environment of the youth [i.e. 4, 5]. Still, these treatments are reimbursed by the Dutch social health insurance system and, as such, part of the health care sector. Therefore, like other health care interventions, each intervention needs to demonstrate value for money since it competes for limited funds with other interventions. Efficiency considerations are deemed important in guiding decisions on which treatments to reimburse or initiate. However, given the atypical aim of these systemic interventions, i.e. reducing youth delinquency, an important question is how these types of interventions could demonstrate their efficiency or value for money. The conventional health economic approach of measuring improvements in terms of QALYs may fall short in this context.

Indeed, considering the literature on reducing youth delinquency, it becomes clear that important differences exist between economic evaluations performed in the health care sector and evaluations of crime prevention and treatment programs. It seems that both fields commonly perform sophisticated effect studies, including randomized controlled trials, meta-analyses and systematic reviews [6-10]. Considering economic evaluations of crime prevention and treatment programs the classical cost-benefit analysis is conventionally used

[11, 12]. An extensive cost-benefit evaluation of crime prevention and intervention programs has been performed by Aos and colleagues [2] in the United States. That evaluation was based on a literature review, computation of average effects per treatment program, assignment of a monetary value to the effects and subsequently calculation of a net present value in a cost-benefit model structure. Furthermore, French and colleagues [13, 14], for example, conducted cost-benefit analyses on addiction treatment for substance abusers. These cost benefit analyses were deterministic models [2, 13, 14]. In addition, Aos and colleagues [2] assessed costs and benefits from a taxpayer perspective. In health economic literature, cost-effectiveness analyses are preferably conducted from a societal perspective. Another difference between the two fields is, that in health economics sophisticated methodological guidelines for economic evaluations have been developed, while in the field of criminal justice such guidelines do not (yet) appear to exist. Furthermore, in health economic literature, cost-effectiveness or cost-utility analyses dominate [12]. In the field of crime prevention and treatment, these analyses are limited. Nevertheless, McCollister and colleagues [15-17] and French and colleagues [18] conducted various cost-effectiveness analyses related to substance abuse treatment, where the effectiveness is for example measured as days of re-incarceration [15-17] or as a delinquency score [18]. These studies show clearly the use of state of the art methods developed in the field of health care, applied in the field of crime prevention and treatment. On the other hand, these cost-effectiveness analyses were relatively conventional as parameter uncertainty was not captured in the model and long-term estimates were not taken into account. A common way to assess the cost-effectiveness in health care is the so-called decision analytic model [19, 20]. This approach provides a mathematical structure, synthesizing the evidence on costs and effects in a treated population under a variety of treatment options and makes the uncertainty around estimates visible. An additional advantage of this decision analytic modeling approach is that long-term effects can be modeled, even beyond the duration of the trial. Decision-analytic modeling and in particular inclusion of long-term effects may be especially relevant for interventions aiming to reduce criminal behavior. Several authors suggested that criminal behavior during adulthood tends to be preceded by behavioral disorders during childhood. Berger and Boendermaker [21] stated that serious offenders often have a history of problematic behavior in their early years of life. Kim-Cohen and colleagues [22] mentioned that most mental disorders in adults "...should be reframed as extensions of juvenile disorders". This suggests that systemic interventions for juvenile disorders may reduce future criminal activity later

on in life. Estimates of long-term effects are therefore essential to the analysis of these interventions.

The current study aims to build a probabilistic decision analytic model like common models in health care for assessing interventions primarily aimed at crime prevention and treatment in youth care. In developing the model the following requirements had to be met:

i. The model should be applicable to assess costs and effects of systemic interventions primarily aimed in reducing delinquent behavior;
ii. The initial model should be fairly simple however easy to adjust to sophisticated details (i.e. severity of delinquency);
iii. The model should be probabilistic, taking uncertainty into account;
iv. The model should be suitable for long-term analysis;

As an illustration an initial assessment of the cost-effectiveness of Functional Family Therapy (FFT) compared to treatment as usual (TAU) is presented. As the aim of the study is the application of the probabilistic decision analytic modeling to interventions aimed at reducing delinquency, the interventions compared could be substituted by other systemic interventions mentioned.

The article is structured as follows. The methods section provides information on the health economic model type and general characteristics of the model. The results section elaborates on the applicability of the decision analytic model and outcome measure to the field of systemic interventions specifying necessary adaptations to the health economic approach based on an initial assessment of cost-effectiveness of FFT. The conclusion relates our findings to the general objective of applying health economic methods to systemic interventions not primarily aimed at improving health.

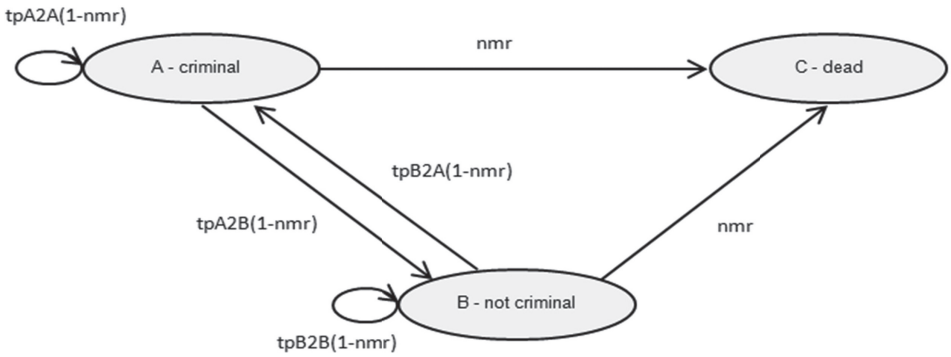
Methods

Model structure

We constructed a probabilistic Markov cohort model [19]. Disease progression in common Markov models is described using transitions between 'states', where a subject can move between states or remain in the current state. The transition rates between states are typically estimated based on short run data. Long-term predictions are made based on repetition of transition cycles and assumptions based on for example literature.

In order to keep the initial model as transparent as possible, a Markov model was constructed consisting of three states, i.e. A - criminal behavior, B - non criminal behavior and C - dead. The model structure is shown in Figure 1. All subjects in our study started in state A, moved to either state B or C or remained in state A and could then move between criminal and non-criminal states. Death acted as the absorbing state. Note that subjects could also remain in their present state (depicted by the u-turns).

Figure 1 | Markov model



nmr = natural mortality rate.
tpA2A = transition probability of staying in state A.
tpA2B = transition probability of moving from state A to state B.
tpB2A = transition probability of moving from state B to state A.
tpB2B = transition probability of staying in state

Outcome measures and model parameters

In order to apply health economic methods meaningfully in the field of crime prevention and treatment, we introduce a new and neutral outcome measure of cost-effectiveness modified for this particular type of intervention: criminal activity free years (CAFYs). The CAFY was defined as a measure of time spent in a dichotomous criminal or non-criminal state. When extensive data is available, criminal activity can e.g. be defined as having had police contacts or committed crimes in the past half year. For the purpose of demonstrating the model functioning and in the absence of extensive clinical data, the criminal state in this study was based on adolescent recidivism derived from clinical trial findings reported by Sexton and Alexander [4]. Transition probabilities differed according to the treatments offered. Treatment costs also differed per treatment type whereas all other costs (Table 1) in the different states were assumed to be independent of the treatment arm but dependent on the state. The cycle length

used in the model was six months. This corresponds to the period common for follow up intervals in clinical trials in the field of crime prevention [6, 7, 23].

Table 1 | Included types of costs

Cost categories	Direct	Indirect
Health care	Medical and mental health care child *(psychologist, psychiatrist, GP, specialist, ER, hospital (day) care, medication, youth welfare agency (bureau jeugdzorg)*, foster home*, residential institution, centre for addiction treatment, social worker)*	
	Medical and mental health care parent *(psychologist, psychiatrist, GP, specialist, foster care*, center for addiction treatment, social worker)*	
Outside health care	Travel expenses (incl. parking)	Productivity losses parent *(absence from work, inefficiency at work)*
	Time spent by child on exercises as part of therapy*	Informal care/ support child *(community centre/ church/ mosque/ association, care/support by family or acquaintances)*
	Time spent by parent on exercises as part of therapy*	Criminal justice system child *(Council of child protection, Bureau Halt*, Police, Lawyer, Court, Incarceration costs)*
		Informal care/ support parent *(community centre/ church/ mosque/ association)*

* Included until age 30

In the developed model two treatment alternatives were compared. To provide an example of a cost-effectiveness analysis of systemic interventions, a group receiving FFT therapy and a comparison group receiving TAU were evaluated. TAU refers to a comparable treatment, which delinquent youth would have received if they had not received FFT. As institutions offer diverse types of alternative therapies to FFT, TAU may differ between the different institutions. In one institution TAU may be MST, while another institution may offer Cognitive Behavioral Therapy (CBT) as an alternative to FFT. In our illustration subjects could not switch between FFT and TAU.

For an extensive comparison between two systemic interventions, the model should include several types of cost categories. Table 1 depicts the common cost categories in health economic evaluations; direct and indirect costs inside

and outside the health care system adapted to the field of crime. The included types of costs are derived from a combination of the costs commonly included in health economic evaluations and literature on cost of crime [24]. These costs not only pertain to costs incurred by the delinquent juvenile, e.g. costs due to criminal activities or treatment, but also to costs falling on family, caregivers and the society as a whole. For reasons of comparability with other interventions in health care, the model included all relevant societal costs in accordance with the Dutch manual for costing in economic evaluations [25].

Discount rates for future costs and effects were set consistent with guidelines for economic evaluations in the Netherlands [26]. (Note that differential discounting is required in the Netherlands to account for the growth in the value of health over time. See for example Brouwer and colleagues [27] for the rationale behind this. Therefore, by using these rates it was implicitly assumed here, that the value of a criminal activity free year (CAFY) will also increase over time, comparable to the rate of a QALY.)

Data analytic procedures: Cost-effectiveness and scenario analyses

In effect studies, uncertainty is generally represented as a confidence interval, i.e. the magnitude of uncertainty is expressed in standard deviations of the measurement error. This assumes that all relevant uncertainty is measurable in a single outcome measure, and that the distribution of the measurement error is reasonably normal. As both assumptions do not apply in typical health economic evaluations, normal t-tests and other parametric statistics are not particularly useful in health economic modeling. Instead, probabilistic analysis was conducted to take the uncertainty of the model parameters into account. In this analysis uncertainty was simulated by running the Markov model several times using a large cohort of subjects, each time with slightly different parameter values. These values were obtained by randomly sampling from each of the parameter distributions, i.e. gamma distributions for costs, and Dirichlet distributions for transition parameters [19]. One thousand Monte Carlo simulations were performed. In each simulation a random draw from the parameter distributions was taken, which creates a unique set of cost and effect parameters. The expected costs and effects were then calculated and could be plotted on a cost-effectiveness plane. Four additional scenarios were run to demonstrate model behavior under different assumptions. As the transition probabilities constitute important model parameters, a scenario was created in which probabilities for both interventions were equal. Subsequently, the

intervention costs are important parameters, since systemic interventions are concerned to be relatively costly [2]. From a societal perspective, family costs are assumed to be important, therefore it was investigated how exclusion of these costs would influence the results in the third scenario.

Results

The resulting health economic model for systemic interventions showed that modelling an intervention with a primary aim of decreasing delinquency was feasible. Based on the illustrative comparison of FFT versus TAU, costs and effects could be expressed in costs per CAFY. This section elaborates on the specific characteristics of the resulting decision analytic model. Obviously, the combination using different sources for the inputs of a model is certainly not without problems, but we stress that the emphasis here was on building an illustrative model and demonstrating the model functioning.

Model structure

Estimates of long-term effects were essential to the analysis and were taken into account in the current model. This required some (informed) assumption regarding the endurance of effects of treatment also taking into account the influence that reaching a certain age or experiencing certain life events may have on criminal behaviour [28]. For the current model, information on these parameters was taken from the literature. Moffitt [29] roughly suggested that after adolescence or at approximately age 30 subjects who are criminal during their entire life, life-course-persistent offenders, will remain criminal and subjects who only show criminal behaviour during their adolescence, so-called adolescence-limited offenders, will have returned to non criminal behaviour. This implies a stable state of criminal activity among individuals of age 30 and older. To illustrate the option of incorporating earlier theory and evidence on the development of offending and antisocial behaviour we integrated parts of the long-term stabilising effects described by Moffitt [29] into the current model framework. This effect is implemented in the model by extending the effectiveness of the treatment till the age of 30 years. Consequently youth remain in their current state after that age. Thus after reaching the age of 30, youth reach a stable state in their criminal behaviour, which means the transition probabilities in the model are from then on defined by mortality rates only. The time horizon of the model is 50 years.

To illustrate how long-term effects may influence model results, Figure 2 and Figure 3 present the percentage of youth in each model state over the time horizon of the model, for FFT and TAU respectively. Figure 2 and Figure 3 demonstrate that a stable state is already reached after about 1 year, which implies that the actual impact of the incorporation of a stabilising effect, based on the theory of Moffit [29] is minor in this model. However, as the model results are only as good as the available input used to fill the model, the current results only illustrate how long-term effects could be included in the present model, as it is mainly based on assumptions made and empirical data are lacking.

Figure 2 | Percentage of youth in model states over time for FFT

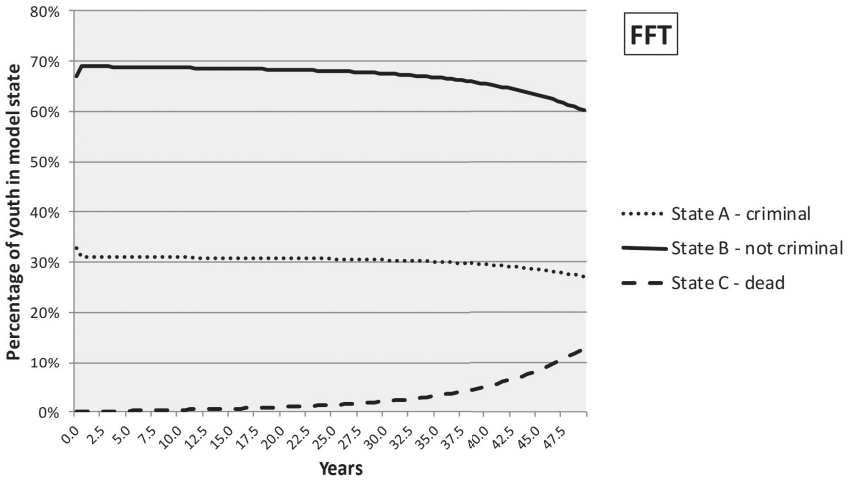
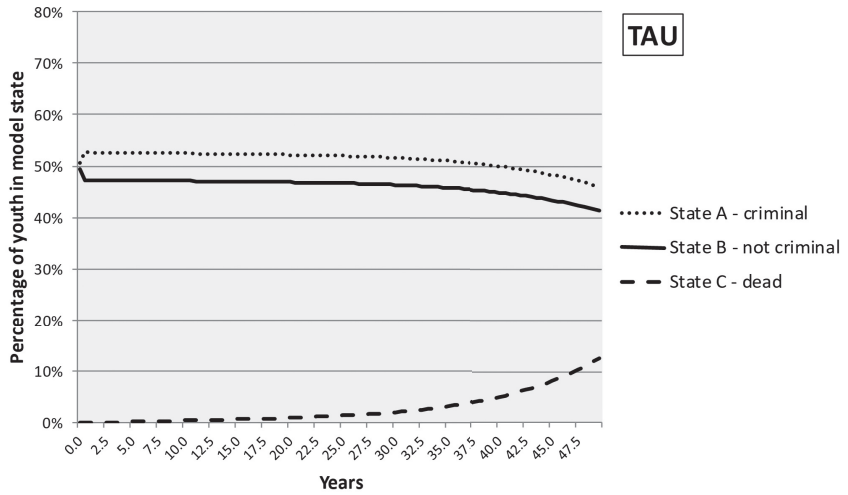


Figure 3 | Percentage of youth in model states over time for TAU



Outcome measure: CAFY

In health economic evaluations, cost-effectiveness is most commonly estimated in cost per quality-adjusted life year (QALY). However, as the predominant effect of behavioral interventions for criminal youths is the reduction of criminal activity [30] and thus is not directly or exclusively linked to physical health and absence of disease, the effect measure QALY seems inadequate to capture the full benefit of interventions in adolescent mental health [31]. Therefore, a different outcome measure that sufficiently captures the goals of crime prevention and treatment was required. Considering the societal perspective of the policymaker, a broad outcome measure, directly linked to the goal of a reduction in criminal activity, was chosen. As a first step in this context, we chose the outcome measure of criminal activity free years (CAFYs), which can be used to determine the (incremental) costs per CAFY, i.e. the costs per criminal activity free year. Using CAFYs as the effect measure enables decisions based on a non-monetary value that is comparable between interventions and that properly reflects the goals of the Ministry of Justice while fitting into the health economic modelling approach. Existing examples of an effectiveness measure that resembles the use of the CAFY measure, is the use of days re-incarcerated [17].

As the model has two states defined as either being criminal or not being criminal, the transition from state A, criminal, to state B, not criminal, represents the rate of not being criminal after treatment. The transition of state B to state A on the other hand represents the rate becoming criminal after having been not criminal. It is assumed all youth enter the model as being criminal. The outcome of (incremental) costs per CAFY, was (as a first and rather simplified step) obtained by assigning different costs to individuals according to their current state, criminal state A or non criminal state B. Determining the net present value of the additional costs incurred in state A and state B over the full lifespan of subjects and dividing these by the amount of additional years the individual spends in the non criminal state B during his entire life (compared to TAU) yielded an estimate of incremental costs per CAFY. This process of calculating life-time costs and dividing these by life-time criminal-activity-free years was repeated 1000 times by means of simulation in order to reflect variability in input parameters.

Model parameters: Transition probabilities

Transition probabilities were dependent on the definition of the states reflecting the choice of outcome measure. In the current model, criminal behaviour was

chosen as most relevant outcome measure so that the states were defined as 'criminal' and 'non criminal' and transition probabilities between the states could be retrieved from literature.

Several studies showed the effectiveness of FFT compared to TAU [4, 18, 32-34]. Yet, there is no consistent outcome regarding the effectiveness of FFT in comparison to TAU. The results based on adolescent recidivism derived from clinical trial findings reported by Sexton and Alexander [4] were most applicable and comparable to the formulation of our model parameters and definition of the comparison group. So demonstrating the model, we used the effectiveness rates of that study [4]. As the rate of recidivism based on the clinical trial reported in the study of Sexton and Alexander is 33 percent [4], we assumed this rate could be equal to the transition from state B to state A in the model and is therefore supposed to be equal to 33 percent. As the sum of all transition probabilities related to one state in the model sums up to 100 percent, the transition rate of state B to state B (individuals remaining in the non criminal state) is set at 67% (100% minus 33%). As for illustrative purposes we assumed here that the probability of individuals staying non-criminal (B to B) to be equal to the probability of becoming non-criminal (A to B),the transition from state A to state B, was fixed at 67 percent as well. Again subtracting this transition rate from 100% resulted in a probability of 33% for individuals remaining in the criminal state (A to A). Sexton and Alexander [4] furthermore suggested that "FFT reduces recidivism and/or the onset of offending between 25 and 60 percent more effectively than other programs". As TAU refers to a comparable treatment, we took the average of this range as a reasonable and illustrative estimate of the effectiveness of TAU. The model therefore was constructed under the illustrative assumption that FFT reduces criminal activity 42.5 percent more effectively than TAU.

Transition probabilities were assumed to be fixed over the years, as no further long term effectiveness is known yet.

Model parameters: Costs

To fill in the cost parameters in the model the costs in the criminal state were retrieved from an ongoing trial of FFT [35]. The volumes of costs in the non criminal state were derived from scaling volumes in the criminal state with a ratio of cost volumes of anti-social versus "normal" youths presented in a UK study on the financial costs of anti-social youths [36]. Unit prices were taken from the Dutch manual for costing in economic evaluations [25]. In absence of Dutch

unit costs, mean treatment costs of the interventions compared were derived from American costs presented in the study of Aos and colleagues [2]. These costs are not related to the states but depend on the intervention a youth received.

Cost-effectiveness

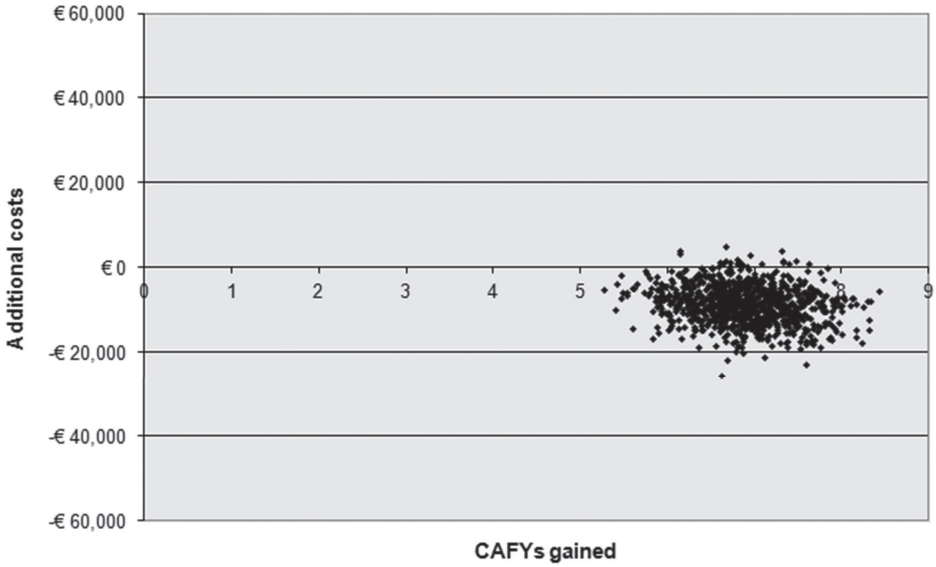
As the comparison of FFT with TAU in the current model is illustrative, the model results solely fulfill this objective. These illustrative cost-effectiveness results from the model point towards lower costs of FFT when compared to TAU. Taking the mean from the stochastic results, the number of CAFYs for FFT exceeds the number of CAFYs for TAU by 6.88 and the costs of FFT appear lower than TAU with incremental cost savings of 8,577EUR (Table 2), positioning the intervention in the South East quadrant of the cost-effectiveness plane (Figure 4). Incremental cost-effectiveness from the illustrative model data expressed in costs per CAFY amounts to cost savings of 1,246 EUR/CAFY. These exemplifying results suggest that FFT produces better effects at lower cost when compared to TAU.

Table 2 | Scenario analyses

	CAFY's gained	Cost savings
Base case	6.88	8,577
Scenario 1: transition rate FFT=TAU	-0.02	-718
Scenario 2: TC FFT = TC TAU	6.85	9,112
Scenario 3: excl. family costs	6.88	6,307

FFT = Functional Family Therapy
TAU = Treatment As Usual
TC = Treatment Costs

Figure 4 | Cost-effectiveness results - Base case analysis



Scenario analyses

Scenario analysis can reveal how the results change if certain parameters are changed. The scenario analysis indicates that the model is particularly sensitive to changes in transition rates whereas the results appear rather robust to changes in other input parameters (Table 2). When transition rates of TAU and FFT are assumed equal (Table 2, Scenario 1), cost savings and CAFY gains entirely vanish. Simulation then results, on average, in an incremental effect of zero and negligible differences in costs between the interventions. The results of the model thus appear to strongly depend on accurate estimates of transition probabilities. Variation in intervention costs does not yield significant differences in costs or effects (Table 2, Scenario 2), whereas exclusion of family costs not only results in a decrease in cost savings but also decreases the variance of the incremental costs (Table 2, Scenario 3).

Discussion and Conclusions

This study created a framework for the evaluation of interventions aimed at reducing criminal activity in delinquent youth. A probabilistic Markov model approach was constructed allowing the assessment of the incremental cost-effectiveness of two systemic interventions. For illustrative purposes, the interventions considered were FFT and TAU. As the comparison of FFT with TAU in the current model is solely an example to demonstrate model functioning,

the model results are illustrative in absence of empirical data. As a first step to come to suitable outcome measures in this field, we introduced the outcome measure of Criminal Activity Free Years (CAFY) in a probabilistic decision analytic model. The presented methodology may provide a basis for further development of the model and outcome measures and, ultimately, decision-making by both Ministries of Justice and, in particular, Health. Policymakers may compare cost and effects between different types of interventions aiming to reduce delinquency among youth.

An advantage of using decision analytic models is that this approach enables calculation of hypothetical scenarios. Hence, questions of policymakers, for example on differences in cost-effectiveness within subgroups of youth or on the optimal age for intervention may be answered. Moreover, the decision uncertainty is represented in the model results by taking into account the uncertainty surrounding the input parameters of the model. The current study showed that it was feasible to apply health economic methodology to assess interventions aimed at reducing delinquency rates. The approach was developed to be consistent with health economic guidelines. To our knowledge, this was the first economic evaluation using decision-analytic modelling in the evaluation of systemic interventions for crime prevention and treatment.

However, a number of important questions remain. First of all, the outcome measure presented here is clearly sector-specific. While this enables choosing between interventions with similar aims, it does not directly allow comparisons with other interventions. This problem is not unique for this context. For instance, interventions in elderly care or social care may not be primarily aimed at producing health as well. Outcome measures such as the OPUS and ICECAP have been proposed as better capturing the benefits of such care [37, 38]. This does raise the question, however, of how to trade-off between interventions when their aim is not similar and when different outcome measures were used to assess cost-effectiveness. This seems to be an important area for future research.

Secondly, we proposed the measure of CAFY as a first step to demonstrate how interventions aimed to reduce delinquency could be evaluated within a probabilistic decision model. If such interventions were to be evaluated more systematically using methodology like the one presented here, clearly, the outcome measure deserves more attention. The outcome measure of the CAFY is a very simple and crude one. One could compare it to 'natural units' used in

cost-effectiveness analysis like gained life years and event free life years. An important problem with these measures and the CAFY is that they do not reflect the seriousness of the events (e.g. living in a poor or good health state or, in this case, engaging in many and severe criminal activities or a few minor felonies). However, the definition of criminal activity free could be based on different measures, like the number of police contacts or youth self-report of committed crimes. Since not all committed crime, irrespective of the seriousness of the crime, is reported to the police, the difference in definition could give different effectiveness and cost-effectiveness results. Preference weighted measures (like the QALY) would be preferred in this context. Such measures could add a weight to different types of criminal activities and be more comprehensive in terms of the benefits they include (which could even entail a mix of health and crime-related outcomes).

Reducing delinquent behavior is an important outcome of systemic interventions, but multiple other outcomes may be relevant as well, among which for example the ability to live at home after treatment, school attendance or family functioning [31, 39]. As these multiple outcomes are not considered in the current model, it could be valuable to extend the model or broaden the outcome measure.

Before further use, the model would require improvement, since our analysis had a number of limitations. First, the model was limited to three states. Although a model is always a simplification of reality, and the current model even was an illustration, it should be investigated whether three states are sufficient to provide reasonable estimations of reality. Secondly, the states used now were dichotomous (criminal or non-criminal behaviour). The severity of criminal offenses is likely to be important as well, also as a predictor of future criminal activity [28]. The frequency or the types of crime could be an important differentiating factor to discriminate more detailed states [28]. Using more differentiated states would therefore add validity to the model. However, a necessary condition for the formulation of a more complex model is the availability of more and detailed trial data. Third, an individual's history of offenses could be used to predict future behaviour and, thus, it may be useful to relax the 'memoryless' feature of the Markov model [19]. This feature encompasses that once a subject has moved from one state to another, the Markov model will have 'no memory' regarding which state the subject has come from or the timing of that transition. Using the history of earlier offences in the model could also improve the resulting estimates. The incorporation of long-term effects in the model was based on the

coarse assumption individuals reach a stable state of criminal behaviour after an age of 30 [29]. However, the impact of using this theory in the current model was minor. In future research one could consider incorporating other relevant theories like the one used here [29] to improve long-term effect modelling. Various other theories and studies about the development of offending and antisocial behaviour exist [28], that could be used to incorporate long-term effects into the model. For example, Sampson and Laub [40] suggest that offending depends on the strength of bonding to society, like bonding to family, peers, school and social institutions. In addition, an early age of onset predicts a relatively long criminal career [11, 28] and several risk factors for the early onset of offending are acknowledged [28]. Besides using studies like those mentioned, a stabilising effect could be modelled more smoothly over time or could be based on empirical, long-term follow-up data to add more detail to modelling long-term effects. Furthermore, Value of Information (VoI) analyses should explore the additional value of further research to characterize the uncertainty of the model inputs, including long-term effects [19]. Fourth, the cost parameters in the model are depicted from a combination of costs used in health economic evaluations and literature on cost of crime. However, victim costs and intangible costs, which include direct economic losses of the victims and indirect losses suffered by these victims, respectively, are not taken into account [41]. Addition of these costs could be of value. Finally, model parameters were solely based on the limited evidence base of available literature and where retrieved out of different literature sources. Ideally, these parameters would be retrieved from more comprehensive empirical data. For example, the transition probabilities could be linked to the presence or absence of police contacts, contacts with judicial institutions or committed crimes. Availability of additional data can refine the input data of the model and increase the validity of the model structure and the accuracy of the results.

Concluding, we used the methods commonly employed in health economic evaluations to create a framework for determining the value for money of interventions targeted at reducing youth delinquency. The results are encouraging, but important further steps still need to be taken. A first next step may be the collection of empirical data to test the presented methodology. We further suggest the construction of a multidimensional outcome measure that enables researchers to capture the multiple dimensions of the treatment goals, in a preference-weighted manner. A final matter that deserves attention is the value we assign to outcomes such as reduced delinquency. Calculating cost-

effectiveness is especially useful when the results can be judged against some 'threshold' value. What this should be in this context remains unclear as yet.

References

[1] Welsh BC, Loeber R, Stevens BR, Stouthamer-Loeber M, Cohen MA, Farrington DP. Cost of juvenile crime in urban areas: A longitudinal perspective. Youth Violence Juvenile Justice. 2008;6:3-27.

[2] Aos S, Lieb R, Mayfield J, Miller M, Pennucci A. Benefits and costs of prevention and early intervention programs for youth. Olympia: Washington State Institute for Public Policy; 2004.

[3] Ministry of Justice. Afstemming van gedragsinterventies voor jeugdige delinquenten Programma Aanpak Jeugdcriminaliteit [Adjustment of behavioural interventions for youth delinquents: Program of procedure youth delinquency]. The Hague, the Netherlands: Ministry of Justice; 2008.

[4] Sexton TL, Alexander JF. Functional Family Therapy. Office of Juvenile Justice and Delinquency Prevention, Juvenile Justice Bulletin. 2000;1:1-7.

[5] Hendriks V, van der Schee E, Blanken P. Treatment of adolescents with a cannabis use disorder: Main findings of a randomized controlled trial comparing multidimensional family therapy and cognitive behavioral therapy in the Netherlands. Drug and Alcohol Dependence. 2011;119:64-71.

[6] Glisson C, Schoenwald SK, Hemmelgarn A, Green P, Dukes D, Armstrong KS, et al. Randomized trial of MST and ARC in a two-level evidence-based treatment implementation strategy. Journal of Consulting and Clinical Psychology. 2010;78:537-550.

[7] Henderson CE, Rowe CL, Dakof GA, Hawes SW, Liddle HA. Parenting practices as mediators of treatment effects in an early-intervention trial of multidimensional family therapy. The American Journal of Drug and Alcohol Abuse. 2009;35:220-226.

[8] King S, Griffin S, Hodges Z, Weatherly H, Asseburg C, Richardson G, et al. A systematic review and economic model of the effectiveness and cost-effectiveness of methylphenidate, dexamfetamine and atomoxetine for the treatment of attention deficit hyperactivity disorder in children and adolescents. Health Technology Assessment. 2006;10:1-146.

[9] Teuffel O, Kuster SP, Hunger SP, Conter V, Hitzler J, Ethier MC, et al. Dexamethasone versus prednisone for induction therapy in childhood acute lymphoblastic leukemia: A systematic review and meta-analysis. Leukemia. 2011;25:1232-1238.

[10] Ttofi MM, Farrington DP, Losel F, Loeber R. The predictive efficiency of school bullying versus later offending: A systematic/meta-analytic review of longitudinal studies. Criminal Behaviour and Mental Health. 2011;21:80-89.

[11] Loeber R, Farrington DP. Young children who commit crime: Epidemiology, developmental origins, risk factors, early interventions, and policy implications. Development and Psychopathology, 2000;12:737-762.

[12] Soeteman DI, Busschbach JJV. Cost-benefit and cost-effectiveness of prevention and treatment. In R. Loeber, W. Slot, P. van der Laan & M. Hoeve (Eds.), Tomorrow's Criminals: The Development of Child Delinquency and Effective Interventions (pp. 215-226). Surrey: Ashgate; 2008.

[13] French MT, McCollister KE, Cacciola J, Durell J, Stephens, RL. Benefit-cost analysis of addiction treatment in Arkansas: specialty and standard residential programs for pregnant and parenting women. Substance Abuse. 2002a;23:31-51.

[14] French MT, McCollister KE, Sacks S, McKendrick K, De Leon G. Benefit-cost analysis of a modified therapeutic community for mentally ill chemical abusers. Evaluation and Program Planning. 2002b;25:137-148.

[15] McCollister KE, French MT, Inciardi JA, Butzin CA, Martin SS, Hooper RM. Post-release substance abuse treatment for criminal offenders: A cost-effectiveness analysis. Journal of Quantitative Criminology. 2003a;19:389-407.

[16] McCollister KE, French MT, Prendergast M, Wexler H, Sacks S, Hall E. Is in-prison treatment enough? A cost-effectiveness analysis of prison-based treatment and aftercare services for substance-abusing offenders. Law and Policy. 2003b;25:63-82.

[17] McCollister KE, French MT, Prendergast ML, Hall E, Sacks S. Long-term cost effectiveness of addiction treatment for criminal offenders. Justice Quarterly. 2004;21:659-679.

[18] French MT, Zavala SK, McCollister KE, Waldron HB, Turner CW, Ozechowski TJ. Cost-effectiveness analysis of four interventions for adolescents with a substance use disorder. Journal of Substance Abuse Treatment. 2008;34:272-281.

[19] Briggs A, Claxton K, Sculpher M. Decision Modelling for Health Economic Evaluation. New York: Oxford University Press; 2006.

[20] Drummond MF, Sculpher MJ, Torrance GW, O'Brien BJ, Stoddart GL. Methods for the economic evaluation of health care programmes (Third edition). Oxford, USA: Oxford University Press; 2005.

[21] Berger M, Boendermaker L. Multisysteembehandeling in Nederland - Voorstel voor de introductie van MST [Multisystemic treatment in the Netherlands - Proposal for the introduction of MST]. Utrecht, the Netherlands: Nederlands Instituut voor Zorg en Welzijn /NIZW Jeugd; 2003.

[22] Kim-Cohen J, Caspi A, Moffitt TE, Harrington H, Milne BJ, Poulton R. Prior juvenile diagnoses in adults with mental disorder: developmental follow-back of a prospective-longitudinal cohort. Archives of General Psychiatry. 2003;60:709-717.

[23] Hogue A, Henderson CE, Dauber S, Barajas PC, Fried A, Liddle HA. Treatment adherence, competence, and outcome in individual and family therapy for adolescent behavior problems. Journal of Consulting and Clinical Psychology. 2008;76:544-555.

[24] Cohen MA. The costs of crime and justice. New York: Routledge-Taylor Francis Group; 2005.

[25] Hakkaart L, Tan SS, Bouwmans CAM. Dutch costing manual for health care. Amstelveen, the Netherlands: The Healthcare Insurance Board (CVZ); 2010.

[26] The Health Care Insurance Board. Guidelines for Pharmacoeconomic Research. Amstelveen, the Netherlands: The Healthcare Insurance Board (CVZ); 2006.

[27] Brouwer WB, Niessen LW, Postma MJ, Rutten FF. Need for differential discounting of costs and health effects in cost effectiveness analyses. BMJ. 2005;331:446-448.

[28] Farrington DP. Developmental and life-course criminology: Key theoretical and empirical issues - The 2002 Sutherland Award Address. Criminology. 2003;41:221-255.

[29] Moffitt TE. Adolescence-limited and life-course-persistent antisocial behavior: a developmental taxonomy. Psychological Review. 1993;100:674-701.

[30] Erkenningscommissie Gedragsinterventies Jeugd. Committee of approval in youth care. Retrieved from http://www.rijksoverheid.nl/onderwerpen/recidive/erkenningscommissie-gedragsinterventies. Accessed in 2011.

[31] Sindelar JL, Jofre-Bonet M, French MT, McLellan AT. Cost-effectiveness analysis of addiction treatment: Paradoxes of multiple outcomes. Drug and Alcohol Dependence. 2004;73:41-50.

[32] Gordon DA, Graves K, Arbuthnot J. The effect of functional family-therapy for delinquents on adult criminal behavior. Criminal Justice and Behavior. 1995;22:60-73.

[33] Sexton TL, Alexander JF. Functional family therapy for at-risk adolescents and their families (chapter 6). In F. W. Kaslow & T. Patterson (Eds.), Comprehensive handbook of psychotherapy (Volume 2) (pp. 117-140). New York: John Wiley & Sons, Inc; 2002.

[34] Sexton T, Turner CW. The effectiveness of functional family therapy for youth with behavioral problems in a community practice setting. Journal of Family Psychology. 2010;24:339-348.

[35] ZonMw. ZonMW programma zorg voor jeugd: weten wat werkt! [ZonMW program for youth]. The Hague, the Netherlands: ZonMW; 2008.

[36] Scott S, Knapp M, Henderson J, Maughan B. Financial cost of social exclusion: follow up study of antisocial children into adulthood. BMJ. 2001;323:191-194.

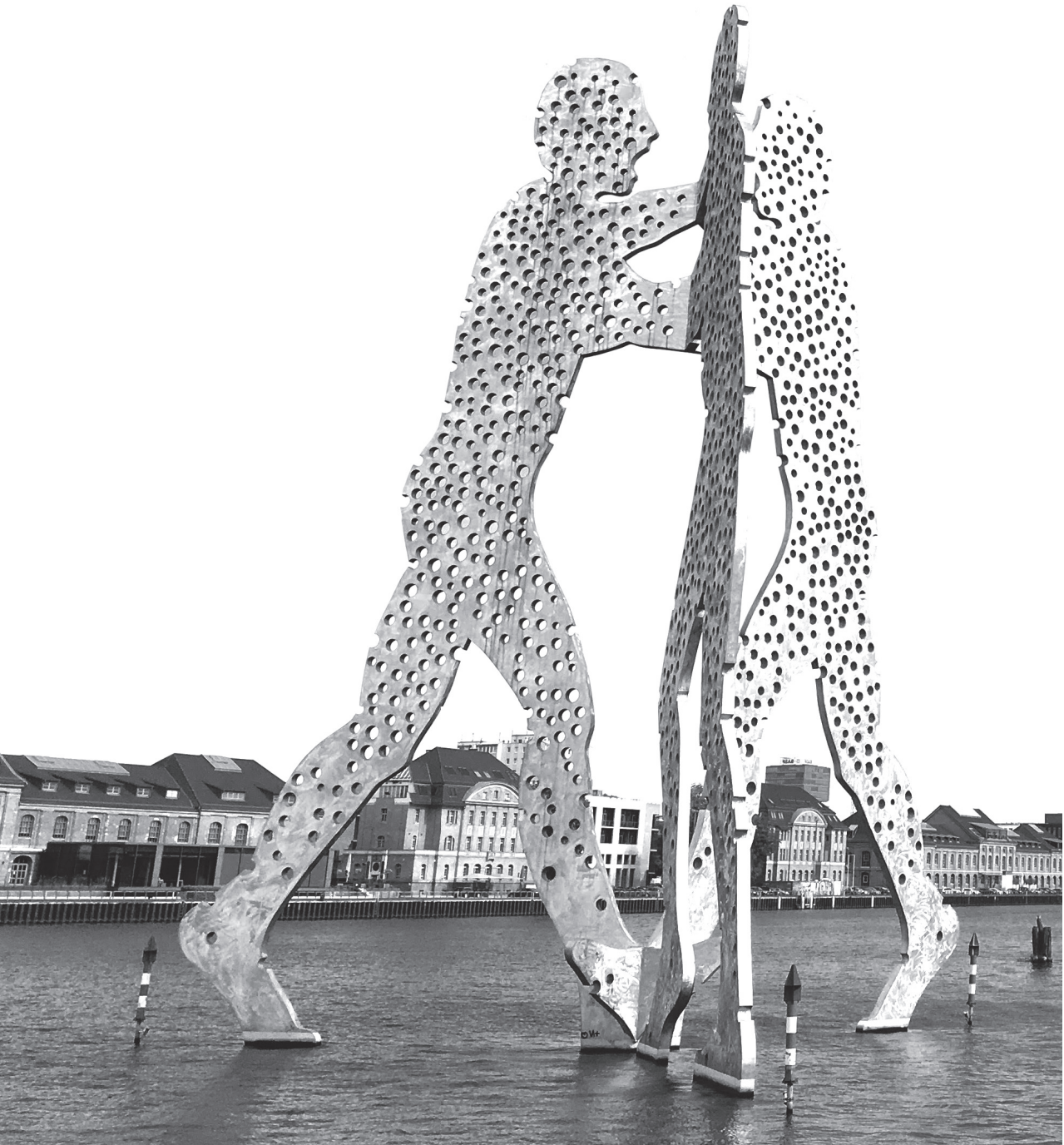
[37] Ryan M, Netten A, Skatun D, Smith P. Using discrete choice experiments to estimate a preference-based measure of outcome--an application to social care for older people. Journal of Health Economics. 2006;25:927-944.

[38] Coast J, Flynn TN, Natarajan L, Sproston K, Lewis J, Louviere JJ et al. Valuing the ICECAP capability index for older people. Social science and Medicine. 2008;67:874-882.

[39] Henggeler SW. Multisystemic therapy: An overview of clinical procedures, outcomes, and policy implications. Child and Adolescent Mental Health. 1999;4:2-10.

[40] Sampson RJ, Laub JH. Crime in the making: Pathways and turning points through life. Cambridge, MA: Harvard University Press; 1993.

[41] McCollister KE, French MT, Fang H. The cost of crime to society: New crime-specific estimates for policy and program evaluation. Drug and Alcohol Dependence. 2010;108:98-109.



Chapter 5

Value of Information Analysis Applied to Systemic Interventions Aimed to Reduce Juvenile Delinquency

Based on van Eeren, H., Schawo, S., Hakkaart-van Roijen, L., Busschbach, J.

PlosOne. 2015; 10(7): e0131255.

Introduction

In order to guide policy decisions, it would be helpful to know the cost-effectiveness of interventions aimed at reducing juvenile delinquency. So far, cost-effectiveness analyses have informed an increasing number of reimbursement decisions in mental health-care [1-2]. Accordingly, the number of cost-effectiveness analyses in the field of crime prevention is increasing [2-10].

The inputs in a cost-effectiveness analysis can be uncertain, as available information about the costs and effects of interventions is rarely perfect. As a result, the decision whether or not to reimburse an intervention is marked by uncertainty. When a decision to reimburse an intervention turns out to be incorrect, it could lead to suboptimal interventions being approved. These interventions create costs in terms of foregone benefits and resources [11-15]. Further research may eliminate this uncertainty and optimize the reimbursement decision.

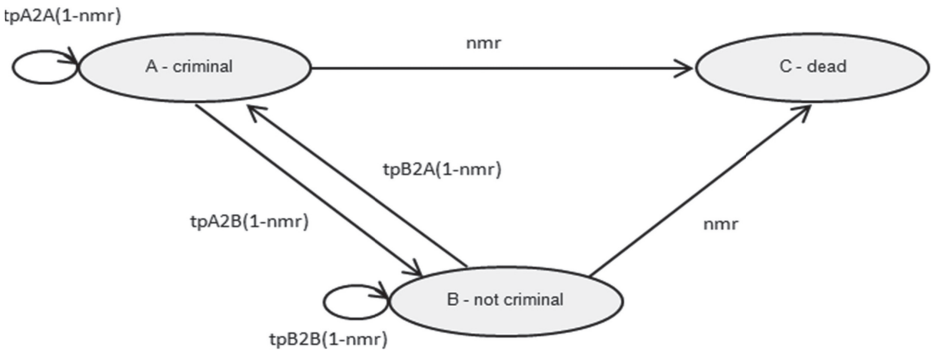
This study aims to estimate the added value of future cost-effectiveness research. This type of analysis is referred to as a 'value of information' analysis and was introduced as part of statistical decision theory [16-17]. It was already applied in other research areas, such as engineering and environmental risk analysis [18], before it was introduced into health technology assessment [11-15,19], where the application of this analysis is now widely adopted, as well as in the field of mental health care [20-21].

A value of information analysis reveals the value of conducting additional research and identifies the type of research that would be most useful. Its results can inform about further research on specific parameters, and more precisely inform the decision about which intervention should be reimbursed [22]. Furthermore, a value of information analysis can be used to prioritize future research, for example by highlighting the merits of certain types of research which might add to the reduction of the parameter uncertainty in cost-effectiveness analysis [15,23-24]. The potential value of further research could then be weighed against the costs of conducting this research in order to determine whether it is worthwhile (i.e. [11-12]).

Because a value of information analysis has not yet been applied in the field of crime prevention, we will present an example of this analysis based on an existing cost-effectiveness model in crime prevention and treatment [25]. We used

two interventions aimed at reducing juvenile delinquency in The Netherlands, for adolescents aged 12-18 years. These interventions can be applied to prevent juvenile delinquency or used to prevent juveniles committing crimes in the future, for example after an adolescent has been punished under the juvenile criminal laws. Juvenile law in The Netherlands applies to adolescents aged 12-17 years [26]. Here, not only the criminal act itself is important, but there is a strong focus on for example the background and moral development of the adolescent [26].

Figure 1 | Markov model.



nmr = natural mortality rate.
tpA2A = transition probability of staying in state A.
tpA2B = transition probability of moving from state A to state B.
tpB2A = transition probability of moving from state B to state A.
tpB2B = transition probability of staying in state

As the present study was set up as an illustration, data was used solely to demonstrate the method. We did not aim to test the superiority of one of the interventions that were used to illustrate the method. Therefore, this article merely presents a demonstration of the relevance of a value of information analysis in the field of crime prevention and treatment. The presented input data and results should be interpreted in this context. We will start with a short summary of the earlier illustrative cost-effectiveness analysis [25], and then introduce and illustrate the value of information analysis.

Methods

Interventions

We compared two interventions aimed at reducing juvenile delinquency. The 'Kursushuis' intervention (translated and referred to as the Course House) consists of a domestic foster home where several adolescents live for about 10

months and professional care is at close hand. The treatment costs and effects were described by Slot et al. [27]. The second intervention is a systemic intervention named Functional Family Therapy (FFT), which lasts about 4 to 6 months. The costs and effects of this intervention were obtained from a multicentre quasi-experimental study in The Netherlands [28]. The Medical Ethical Committee of the VU University Amsterdam approved this study (number 2008/152).

Cost-effectiveness model

The Markov model that was used for the value of information analysis consists of three mutually exclusive model states: A) criminal behavior, B) no criminal behavior, and C) dead [25] (Figure 1). The time horizon of the model was 20 years, with a cycle length of six months [25]. A societal perspective was taken and results were expressed as costs per Criminal Activity Free Year (CAFY) [25].

In line with health economic guidelines [29], the input parameters in the model were threefold. The first group of parameters were the transition probabilities. These reflect the probability that an adolescent transitions through the states. The measure of time an adolescent spends in a non-criminal state is used to estimate a CAFY. Criminal activity was based on the adolescents' self-reported contact with police in connection with him/her having committed one or several crimes; having had no contacts was defined as criminal-activity free and having had one or more contacts as criminally active. Transition probabilities were extrapolated until the age of 30, as we integrated parts of the long-term stabilizing effects described by Moffitt [25, 30]. Dying because of committing crimes was not reflected in the CAFY. Adolescents were assumed to face a risk of death equivalent to the age specific mortality rates in the general population [31]. The second group consisted of costs of health-care use, productivity losses, and other societal costs such as costs of the criminal justice system. Both costs outside health care, and health care costs were included, such as the costs of visiting a psychiatrist or psychologist. As the family system is involved in the interventions provided, we included both the costs of the adolescent and those of one of the parents. The model state costs were fixed over time until the adolescent was 23 years. It was assumed that from that age onwards not all cost categories (such as a family guardian or foster care) would remain relevant. The third group comprised the intervention costs. The costs of one completed FFT treatment were calculated to be approximately €10,900 per adolescent, whereas the Course House was about €37,800 (retrieved from Slot et al. [27]). Both costs were extrapolated to 2013 Euro's accounting for inflation based on the consumer price index [32].

The costs and effects in the model were discounted [33, 34], according to the guidelines for economic evaluations in The Netherlands [29].

To represent the uncertainty of each model parameter, we assigned parameter distributions (S1 table). In a probabilistic analysis, uncertainty was simulated by running the model 10,000 times using a cohort of subjects and each time taking different parameter estimates from the parameter distributions [11-12]. These 10,000 unique sets of parameter values were used to estimate the mean expected cost-effectiveness. For further details on the cost-effectiveness model, we refer to Schawo et al. [25].

Cost-effectiveness analysis

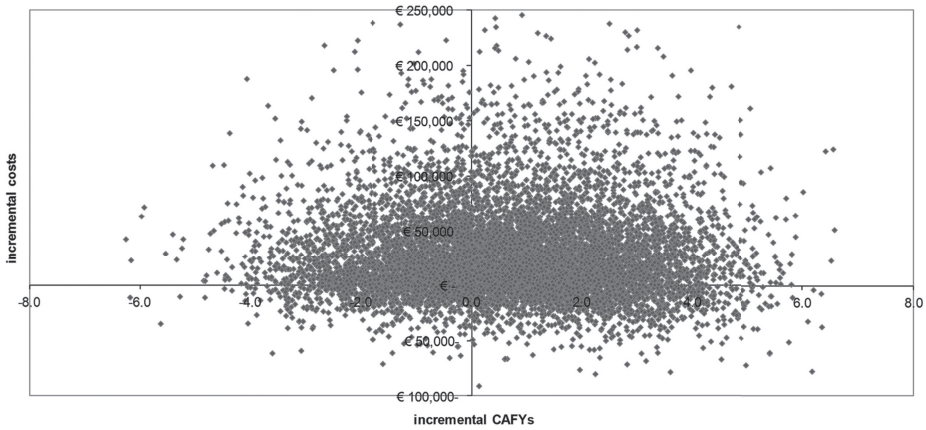
The stochastic model resulted in the relative cost-effectiveness outcomes of the Course House intervention compared with FFT, represented as incremental costs/CAFY (Table 1; Figure 2). It showed that the Course House was more effective than FFT, but also produced higher costs. The cumulative number of CAFYs for the Course House exceeded the number of CAFYs for FFT by 0.7, while the incremental costs of the Course House exceeded those of FFT by €26,800, thereby positioning the intervention in the North East quadrant of the cost-effectiveness plane [35] (Figure 2). The incremental cost-effectiveness ratio (ICER) of the Course House compared with FFT was 39,000 €/CAFY.

Table 1 | Cost-effectiveness results over 20 years[a]

Intervention	Cost	CAFY	ICER[b]	NMB[c]
Course House	€249,000	12.4	€39,000	€641,200
FFT	€222,200	11.7	-	€618,700

[a] The results represented were averaged over the 10,000 simulations run.
[b] The incremental cost-effectiveness ratio (ICER) was calculated as the difference in cost divided by the difference in CAFYs between the Course House and FFT.
[c] The net monetary benefit (NMB) was calculated by multiplying CAFYs by the WTP value of €71,700 per CAFY and subtracting cost. The Course House is cost-effective compared with FFT, because the NMB of the Course House is higher than the NMB of FFT. Due to decimals, the numbers in the table multiplied do not give the exact NMB values represented in this table.

Figure 2 | Incremental cost-effectiveness plane for Course House compared with FFT (10,000 simulations)



Parameter uncertainty

The influence of parameter uncertainty on the model outcomes was shown in the cost-effectiveness-acceptability frontier (CEAF). In the CEAF, the probability of being cost-effective compared to the other intervention is shown for the intervention with the highest expected net monetary benefit (NMB) for a range of societal willingness-to-pay (WTP) values per CAFY, and is therefore cost-effective compared with the alternative intervention [36, 37].

Here, the overall maximum expected net benefit guides the decision on which intervention is cost-effective compared with alternative intervention [36, 37]. The NMB was calculated by multiplying CAFYs by the WTP value per CAFY and subtracting cost [11, 12]. The CEAF is illustrated in the results' section.

Value of information analysis

In the value of information analysis, the parameter uncertainty in the model is monetarized. More precisely, we estimated the value of 'knowing everything': the 'expected value of having perfect information' (EVPI) [12,14]. Having perfect information would eliminate parameter uncertainty and optimize the reimbursement decision. In estimating the 'value of knowing everything', the EVPI places an upper boundary on the value of performing further research [11, 12]. It can be interpreted as the maximum value society 'should' be willing to pay for additional evidence to reduce decision uncertainty around which intervention is preferred and, therefore, inform the reimbursement decision in the future ([11, 12]). The EVPI is computed by first taking the difference between the expected

NMB with perfect information and the expected NMB with current information per simulation. This difference is equal to the expected benefits foregone when making the decision based on current evidence [11, 12]. Comparing the EVPI estimates with the costs of this future research reveals whether further research is worthwhile.

As the value of further information is related to the size of the eligible population of adolescents to be treated, the EVPI was multiplied with the eligible population of adolescents in the population EVPI (pEVPI). About 825 adolescents annually were assumed to be eligible for FFT in The Netherlands. When we discount this number over five years, which is the assumed lifetime of the intervention for which additional research would be useful [11, 29], it resulted in an eligible population of 3,820 adolescents. We assumed that the eligible number of adolescents for the Course House was equal to that for FFT.

In a value of information analysis one could also focus on specific groups of model parameters. To identify the model parameters that contribute to most of the uncertainty and for which future research is the most promising, we estimated the expected value of partial perfect information (EVPPI) [11, 12]. The EVPPI was estimated using the Sheffield Accelerated Value of Information application of Strong et al. [38]. Multiplying the EVPPI values with the eligible population results in the population EVPPI (pEVPPI).

The EVPI and EVPPI not only depend on the uncertainty of the model parameters, but also on the WTP per CAFY. In the absence of a WTP per CAFY in The Netherlands, we used WTP estimates to reduce crime of Cohen et al. [39, 40]. These WTP values per crime indicate the value society wants to pay to prevent one crime, for example €32,200 per burglary (Table 2). Table 2 provides an overview of these estimates, adjusted for inflation and purchasing power parities [41]. Although WTP to prevent one crime is definitely not equal to WTP per CAFY, we used it to illustrate what is meant by WTP in crime prevention and how the concept can be used in a value of information analysis. We hereby implicitly assumed that one crime is committed per year, and thus exactly one crime per year is avoided in a CAFY. We estimated the EVPI and EVPPI for various WTP values, and we chose an average WTP value to illustrate the result in the results section, which was €71,700 (Table 2).

Table 2 | Willingness-to-pay values for crimes (Cohen & Piquero, 2009)

Crime	WTP in 2007 dollars	WTP in 2013 euro's
Murder	$140,000	€128,700
Rape	$290,000	€266,600
Armed robbery	$280,000	€257,400
Robbery	$39,000	€35,900
Aggravated assaults	$85,000	€78,100
Simple assaults	$19,000	€17,4500
Burglary	$35,000	€32,200
Moter vehicle theft	$17,000	€15,600
Larceny	$4,000	€3,700
Druk driving crash	$60,000	€55,200
Arson	$115,000	€105,700
Vandalism	$2,000	€1,800
Fraud	$5,500	€5,100
Other offenses	$1,000	€900
Average	**$140,000**	**€71,700**

The model parameters were grouped into the following ten subsets to indicate the direction of research as a result of the EVPPI analysis: research on 1) transition probabilities for FFT; 2) transition probabilities for the Course House; 3) direct health-care costs of the criminal state; 4) direct health-care costs of the non criminal state; 5) direct non health-care costs related to the criminal state; 6) direct non health-care costs related to the non criminal state; 7) indirect non health-care costs related to the criminal state; 8) indirect non health-care costs related to the non criminal state; 9) intervention costs of FFT; and 10) intervention costs of the Course House.

Results

Model uncertainty

The CEAF shows that FFT had the highest NMB for a WTP ranging from €0 - €39,000 (Figure 3). At a WTP of €39,000, FFT was cost-effective in 49% of the 10,000 model simulations, or a probability of 0.49, whereas the Course House was cost-effective in 51% of the simulations. Above the €39,000 WTP, the Course House had the highest NMB and thus was the optimal intervention. This switching point in the CEAF is where the NMB for FFT is equal to the NMB of the Course House. At this point, the WTP was exactly equal to the ICER value (€39,000 per CAFY).

The CEAF (Figure 3) showed a large error probability. At the WTP of €71,700 the Course House was cost-effective in 57% of the 10,000 model simulations,

which suggests that there is an error probability of 0.43 that could be reduced by collecting additional evidence.

Value of information analysis

In order to know the value of reducing the error probability and to assign a value to additional research, we estimated the EVPI. Table 3 illustrates the EVPI estimation (based on Soeteman et al. [21]). The table shows the generated NMB for each intervention for 6 of the 10,000 simulations, given a WTP value of €71,700 per CAFY. The EVPI was determined as follows: First, we assumed that decision makers have perfect information for each simulation instead of making one single choice over all simulations. For example, for simulation 1 and 2, this would result in the choice for FFT (see Table 3). Second, we determined the choice based on current information. In this case, the Course House had the highest expected NMB (€641,200) over all simulations and hence was the preferred intervention. Finally, we took the difference between the decision based on perfect information per simulation and the optimal choice over all simulations. This difference resulted in the EVPI value or the benefits forgone per simulation. The expectation of all benefits forgone over the 10,000 simulations is the EVPI per adolescent, which is €46,000 at a WTP of €71,700 per CAFY. Perfect information for an individual adolescent was thus valued at €46,000. Multiplying this EVPI value by 3,820 eligible adolescents resulted in a pEVPI of €176 million. This pEVPI value suggests that, at a societal WTP value of €71,700 per CAFY, there is room to reduce the uncertainty in the model by a maximum of €176 million.

Figure 3 | Cost-effectiveness Acceptability Curve (CEAC)

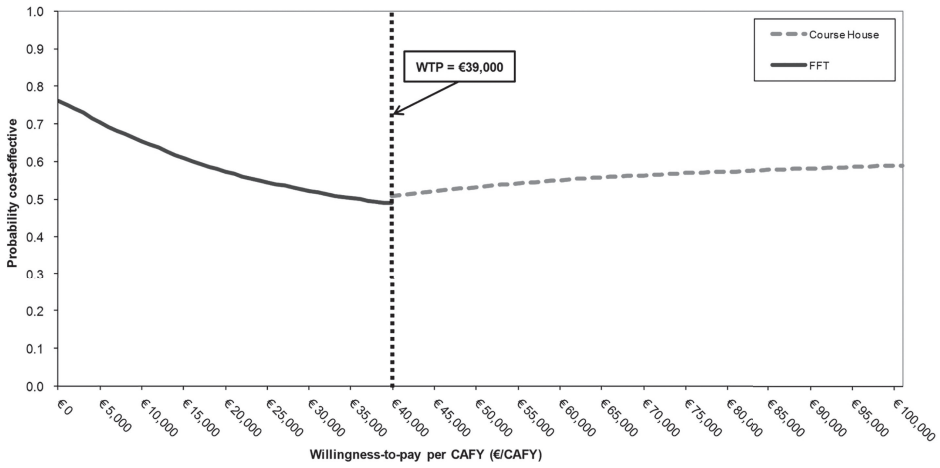


Table 3 | Calculation of expected value of perfect information (EVPI) for individual adolescent

Simulation		Net monetary benefits[a]		Maximum net benefit	Benefits forgone
		Course House	FFT		
	Expectation	***€641,200***	*€618,700*	*€ 687,200*	*€ 46,000*
1		€481,000	**€650,000**	€ 650,000	€ 169,000
2		€553,800	**€710,300**	€ 710,300	€ 156,500
3		€513,800	**€768,000**	€ 768,000	€ 254,200
4		**€717,500**	€562,700	€ 717,500	€ 0
5		€516,200	**€671,000**	€ 671,000	€ 154,800
...	...	...	...	...	...
10,000		**€602,300**	€ 587,200	€ 602,300	€ 0

[a] Net monetary benefit (NMB) was calculated by multiplying CAFYs by the threshold value of € 71,700 per CAFY and subtracting cost.

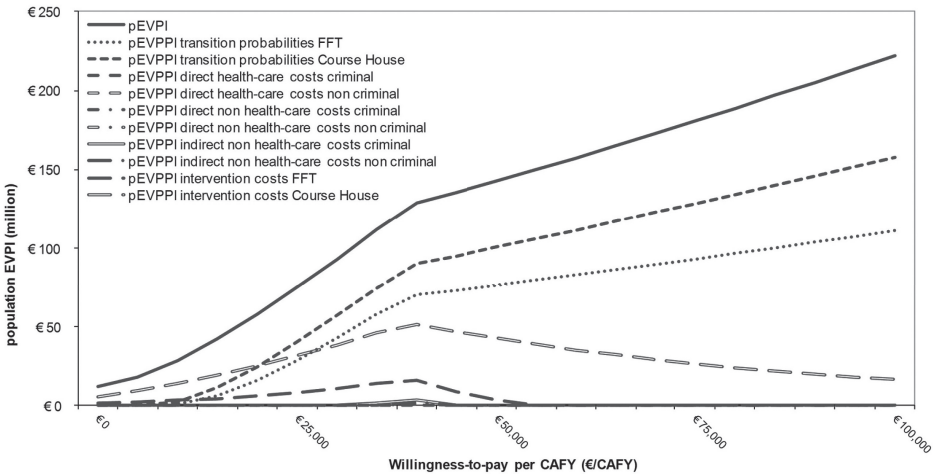
Explanation:
Decision based on current information: Course House.
Decision based on perfect information: bold.
Expected net benefit with current information: € 641,200.
Expected net benefit with perfect information: € 687,200.
Expected value of perfect information (EVPI): € 687,200 - € 641,200 = € 46,000.

Perfect information can be valued at different WTP values. The extent of the monetarized uncertainty surrounding the decision for a range of WTP values is represented in the pEVPI curve. Figure 4 presents the pEVPI curve for an eligible population of 3,820 adolescents. As an example we consider the point where research costs society €50 million (i.e. the pEVPI value at the y-axis in Figure 4). At this point further research would potentially be cost-effective if society were willing to pay more than €17,600 per CAFY (i.e. the value at the x-axis, if the pEVPI is €50 million). At lower values of the WTP per CAFY, the benefits of further research cannot offset the costs [11, 42]. At a WTP of €39,000 per CAFY, the pEVPI shows a local maximum of €127 million. At this point, the parameter uncertainty in the model is the highest and thus decision uncertainty is highest, as already shown in the CEAF curve (Figure 3).

Perfect information of subsets of parameters was valued in the pEVPPI. This pEVPPI was estimated for a range of WTP values (Figure 4). At the illustrative WTP value of €71,700 per CAFY, future research would be most valuable for three subsets of parameters: the transition probabilities and the intervention costs of the Course House and the transition probabilities of FFT (see Figure 4). The pEVPPI of the transition probabilities of the Course House was €125 million (€32,700 per adolescent), and the pEVPPI for the transition probabilities of FFT was €91 million (€23,800 per adolescent). The pEVPPI for the intervention costs of the Course House was €28 million (€7,400 per adolescent). The pEVPPIs for

the direct non health-care costs of the criminal state and the non criminal state were respectively €8,400 and €43,300 (respectively €2 and €11 per adolescent). The pEVPPIs for the other parameter groups were all estimated to be zero (Figure 4), meaning there was no potential value of further research into these parameters. Given a WTP €71,700 per CAFY further research for these parameters would not reduce decision uncertainty. The EVPI and EVPPI values depend highly on the WTP value per CAFY, as can be seen in Figure 4 and Table S2. At a WTP of for example €40,000, there was indeed potential value of further research into all model states costs. Note that due to the interactions within the model structure, the pEVPPI for the groups of parameters do not sum up to the overall pEVPI for the model (see Figure 4 [11, 42].

Figure 4 | Cost-effectiveness Acceptability Frontier (CEAF)



Discussion

While cost-effectiveness analyses are increasingly being used in the field of crime prevention, the value of further research has not yet been estimated for comparison between interventions aimed at reducing juvenile delinquency. An earlier developed cost-effectiveness model was used to estimate this value of further research. This study demonstrated that it was feasible to estimate this value of conducting further research in this context, using a value of information framework common in health economic evaluations. The results can be interpreted as similar to cost/QALY (quality-adjusted life year) studies in health care evaluation.

In this value of information analysis, the results indicated the parameters for which further research was valuable. Our findings show particular uncertainty in three groups of parameters: the transition probabilities of the Course House and of FFT, and to a lesser extent, the intervention costs of the Course House and the direct non health-care costs in both model states. Performing additional research in the suggested fields can reduce parameter uncertainty, and hence, can reduce decision uncertainty.

Therefore, the results of a value of information analysis can prioritize further research to optimize the final reimbursement decision, thereby increasing the probability that adolescents will be assigned to the intervention that is cost-effective, compared with the alternative. Given this information, future interventions could be reimbursed (or not), and they could also be approved 'only in research' (OIR) (i.e. further research is required before the intervention can be approved) or 'approved with research' (AWR) (i.e. research can be conducted while the intervention is approved) [43, 44]. For example, from this study we can conclude that given a WTP of €40,000 per CAFY, the Course House could be 'approved with research'. The Course House would then be reimbursed while further research would be required, for example on the effectiveness of the Course House. Current practice in adolescent care in The Netherlands illustrates this approval condition: the Dutch Youth Institute identifies effective youth interventions, while still conducting research on the effectiveness of some of these interventions [45]. However, approval might lead to irrecoverable costs when the approval is revised due to subsequent research revealing that the Course House was not as effective as expected. Then, approval 'only in research' might be preferred, because commitment to future costs is avoided until the results of further research are known. Approval might even be dependent on any change in the effective price of an intervention [44, 46].

This study was a first attempt to apply a value of information framework to the field of crime prevention and treatment of juvenile delinquents. Therefore, some considerations should be kept in mind. The value of information analysis estimates the monetary value of eliminating all or part of the parameter uncertainty of the presented model. However, two other sources of uncertainty can influence the results: structural and methodological uncertainty [47, 48]. Structural uncertainty relates to structural aspects of the model [47, 49, 50], such as the conceptual framework or the transitions between the model states [50], and it can lead to different estimated model outcomes (i.e. [51]). This structural

uncertainty is likely to be present in our model. For example, we did not account for the severity of crimes in the model states or the elevated risk of death for adolescents in the criminal state (i.e. [52, 53]). The uncertainty of these aspects was not represented in the current value of information analysis. Future cost-effectiveness models in the field of crime preventions should therefore carefully characterize the structural uncertainty [50], and account for it when possible, for example by parameterization [49, 50] or model averaging [47, 50, 54].

The second additional source of uncertainty is methodological uncertainty, which relates to the analytical method chosen [48, 49]. Our model also represents some methodological uncertainties, such as whether or not to include the costs of crime in the model (i.e. [55]). Here, three methodological uncertainties in our model are discussed in more detail. These uncertainties could be resolved through, for example, formulating guidelines (i.e.[47, 49]) to model cost-effectiveness research in the field of crime prevention.

The first methodological uncertainty concerns the societal perspective used in the model, which means that we included the costs and effects relevant to society. When considering this perspective in health care, the focus is merely on the patient, whereas this will be different in the area of crime prevention and treatment (i.e.[56]). In this study, we already included the direct and indirect costs of one parent, as well as direct non medical costs of the adolescent, such as the costs of contact with the police. Other costs that we did not account for, but are nevertheless relevant in the field of crime prevention are: the effect of the intervention reflected in both costs and effects, such as increased wellbeing and reduced productivity losses (i.e.[56]), in regard to family members (e.g. parents or siblings of the adolescents). Further additional categories are reduced victim costs and increased victim wellbeing, the reduction of the number of out-of-home placements, the reduction of the costs of committed crimes to society, reduced costs of avoided crimes to society and the value of reduced fear of crime (i.e. [39, 56]).

The second methodological uncertainty deals with the WTP value for a CAFY. Although we used the WTP values of Cohen et al. [39, 40] to illustrate the use of WTP in crime prevention, WTP to prevent a crime like burglary is definitely not equal to WTP per CAFY. Therefore, it is important to carefully estimate the WTP value per CAFY. In this study, for example, we could have weighted the WTP values of Cohen et al. [39, 40] by the frequency of the crimes as yearly committed

by the adolescents in this study, or by the number of yearly registered crimes in the Netherlands. For clarity reasons and due to a lack of more detailed data regarding the crimes committed, we chose not to use a weighted WTP value. Furthermore, for the WTP values used, it is not exactly known which components of crime, such as investigation, prosecution, witnesses, legal aid, prevention programs, the costs of victims, and the valuation of fear [39] are included in this valuation [39,57]. Therefore, further research is needed into which categories of costs of crime are included in a WTP value, before determining what society is willing to pay for one CAFY. The cost categories included in the cost-effectiveness model should also be reflected in the WTP value and vice versa. Also, other estimations of the societal WTP might be considered. These could, for example, be based on the cost of crime using a bottom-up approach or a breaking-down approach [58]. These methods take into account only the costs of crime, not the willingness to reduce crime levels.

Third, to estimate a WTP per CAFY, it should be known what type of criminal activity is avoided in a CAFY. The seriousness of the crime, the number of times the crime is committed and the types of criminal activity van also be taken into account in defining criminal activity. Furthermore, it is important to decide on how to measure criminal activity. The CAFY used in our study was based on the adolescents' self-reported contact with the police. However, criminal activity may as well be determined on the basis of police registries [27], contacts with other judicial institutions [6-8], rates of reconviction [3] or a delinquency score [5]. Different definitions of criminal activity can influence the model results. For example, not all committed crimes are recorded in police registrations, while self-reported measures could yield socially desirable answers. Using the CAFY in further research thus requires a clear definition of criminal activity.

A final remark on this analysis concerns the interventions chosen. FFT and the Course House were chosen to illustrate the analysis in the field of crime prevention. The interventions under study, however, however could be replaced by other interventions aimed at reducing juvenile delinquency, such as Multisystemic Therapy, Multidimensional Foster Treatment Care or Multidimensional Family Therapy [45]. Contrary to a broader range of cost-effectiveness studies in the UK and US (i.e. [4, 59]), in The Netherlands, to the best of our knowledge, the cost-effectiveness of such interventions has not yet been investigated, except for a cost-benefit analyses of 'Maatregel Inrichting Stelstelmatige Daders' or a case study into 'Strafrechtelijke Opvang Verslaafden'

[56, 60], which are both aimed at adults. Furthermore, in studying interventions in this field, the context of the interventions under study is highly important. In our illustration, we assumed that in practice, the interventions would be applied completely equivalently. However, preferences for an intervention may influence the choice for a certain intervention in reality, such as earlier experience with an intervention, specific characteristics of an adolescent, or the availability of the intervention itself. In our illustration, FFT may be, for example, used more often to avoid committing crimes, whereas the Course House could be used as an addition to a punishment under juvenile justice law, where the adolescent has already committed a crime. There may then be a higher probability of recidivism if the adolescents already have a history of committing crimes [61]. These non-equivalent baseline situations may influence the measured effectiveness of the intervention. Moreover, the situation after treatment may also be different. This may affect the acceptance of possible or required further care (i.e.[61]), and therefore may influence the final degree of committing crimes in the future. In modeling the cost-effectiveness of interventions in the field of crime prevention, the application of interventions in practice should therefore be taken into account in a cost-effectiveness model, or at least, this should be clarified when modeling the cost-effectiveness of such interventions.

In conclusion, an analysis to estimate the value of performing further research had not yet been conducted in the field of crime prevention. The findings of the current study illustrate how such an analysis might be estimated and interpreted in this field. Future investments in cost-effectiveness research on interventions aimed at reducing juvenile delinquency could use this value of information framework to efficiently conduct further cost-effectiveness research.

References

[1] Evers S, Salvador-Carulla L, Halsteinli V, McDaid D, Group M. Implementing mental health economic evaluation evidence: Building a bridge between theory and practice. Journal of Mental Health 2007;16:223-241.

[2] Knapp M, McDaid D, Evers S, Salvador-Carulla L, Halsteinli V, et al. Cost-effectiveness and mental health (MHEEN Policy briefing). London: MHEEN network; 2008.

[3] Barrett B, Byford S. Costs and outcomes of an intervention programme for offenders with personality disorders. The British Journal of Psychiatry. 2012;200:336-341.

[4] Cary M, Butler S, Baruch G, Hickey N, Byford S. Economic evaluation of Multisystemic Therapy for young people at risk for continuing criminal activity in the UK. PloS ONE. 2013;8:e61070.

[5] French MT, Zavala SK, McCollister KE, Waldron HB, Turner CW, et al. Cost-effectiveness analysis of four interventions for adolescents with a substance use disorder. J Subst Abuse Treat. 2008;34:272-281.

[6] McCollister KE, French MT, Inciardi JA, Butzin CA, Martin SS, et al. Post-release substance abuse treatment for criminal offenders: A cost-effectiveness analysis. Journal of Quantitative Criminology. 2003;19:389-407.

[7] McCollister KE, French MT, Prendergast M, Wexler H, Sacks S, et al. Is in-prison treatment enough? A cost-effectiveness analysis of prison-based treatment and aftercare services for substance-abusing offenders. Law & Pol'y. 2003;25:63-82.

[8] McCollister KE, French MT, Prendergast ML, Hall E, Sacks S. Long-term cost effectiveness of addiction treatment for criminal offenders. Justice Quarterly. 2004;21:659-679.

[9] Romeo R, Byford S, Knapp M. Annotation: Economic evaluations of child and adolescent mental health interventions: A systematic review. Journal of Child Psychology and Psychiatry. 2005;46:919-930.

[10] Soeteman DI, Busschbach JJV. Cost-benefit and cost-effectiveness of prevention and treatment. In: Loeber R, Slot NW, Van der Laan PH, Hoeve M, editors. Tomorrow's criminals: The development of child delinquency and effective interventions. Hampshire: Ashgate Publishing Ltd.; 2008. pp. 215-228.

[11] Briggs A, Claxton K, Sculpher M. Decision Modelling for Health Economic Evaluation. New York: Oxford University Press; 2006.

[12] Claxton K. Exploring uncertainty in cost-effectiveness analysis. Pharmacoeconomics. 2008;26:781-798.

[13] Claxton K, Neumann PJ, Araki S, Weinstein MC. Bayesian value-of-information analysis (An application to a policty model of Alzheimer's disease). Int J Technol Assess Health Care. 2001;17:38-55.

[14] Claxton K, Sculpher MJ, Drummond MF. A rational framework for decision making by the National Institute For Clinical Excellence (NICE). The Lancet. 2002;360:711-715.

[15] Oostenbrink JB, Al MJ, Oppe M, Rutten-van Mölken MPMH. Expected value of perfect information: An empirical example of reducing decision uncertainty by conducting additional research. Value in Health. 2008;11:1070-1080.

[16] Pratt J, Raiffa H, Schlaifer R. Statistical decison theory. Cambridge, MA: MIT Press; 1995.

[17] Raiffa H. Decision analysis: Introductory lectures on choices under uncertainty. New York: Random House; 1968.

[18] Yokota F, Thompson KM. Value of information analysis in environmental health risk management decisions: Past, present, and future. . Risk Anal. 2004;24:635-650.

[19] Claxton K. The irrelevance of inference: a decision-making approach to the stochastic evaluation of health care technologies. J Health Econ. 1999;18:341-364.

[20] Mohseninejad L, Van Baal P, Van den Berg M, Buskens E, Feenstra TL. Value of information analysis from a societal perspective: A case study in prevention of major depression. Value in Health. 2013;16:490-497.

[21] Soeteman DI, Busschbach JJ, Verheul R, Hoomans T, Kim JJ. Cost-effective psychotherapy for personality disorders in the Netherlands: The value of further research and active implementation. Value in Health. 2011;14:229-239.

[22] Myers E, Sanders GD, Ravi D, Matchar D, Havrilesky L, et al. Evaluating the potential use of modeling and value of information analysis for future research prioritization within the evidence-based practice center program (AHRQ Publication No. 11-EHC030-EF). Rockville, MD: Agency for Healthcare Research and Quality; 2011.

[23] Sculpher MJ, Claxton K. Establishing the cost-effectiveness of new pharmaceuticals under conditions of uncertainty - When is there sufficient evidence? Value in Health. 2005;8:433-446.

[24] Carlson JJ, Thariani R, Roth J, Gralow J, Henry NL, et al. Value-of-information analysis within stakeholder-driven research prioritization process in a US setting: An application in cancer genomics. Med Decis Making. 2013;33:463-471.

[25] Schawo SJ, van Eeren H, Soeteman DI, Van der Veldt MCAE, Noom MJ, et al. Framework for modeling the cost-effectiveness of systemic interventions aimed to reduce youth delinquency. Journal of Mental Health Policy and Economics. 2012;15:187-196.

[26] Van der Laan PH. Just Desert and Welfare: Juvenile Justice in the Netherlands. In: Junger-Tas J, Decker SH, editors. International handbook of juvenile justice. New York: Springer; 2006. pp. 145-172.

[27] Slot NW, Jagers JD, Beumer MH. Tien jaar Kursushuis: Ervaringen en follow-up-gegevens. Kind en adolescent. 1992;13:62-71.

[28] Van der Veldt MCAE, Eenshuistra RM, Campbell EE. FFT versterkt: Een evaluatiestudie naar de implementatie en de effecten van Functional Family Therapy in Nederland [in Dutch]. Amsterdam: PI Research; 2011.

[29] The Health Care Insurance B. Guidelines for Pharmacoeconomic Research [in Dutch]. Amstelveen, The Netherlands: CVZ; 2006.

[30] Moffitt TE. Adolescence-limited and life-course-persistent antisocial behavior: A developmental taxonomy. Psychol Rev. 1993;100:674-701.

[31] Netherlands S. Statistics Netherlands: Life tables. 2010.

[32] Statistics N. Statistics Netherlands: Consumer price index. 2012.

[33] Brouwer WB, Van Hout BA, Rutten F. A fair approach to discounting future effects: Taking a societal perspective. J Health Serv Res Policy. 2000;5:114-118.

[34] Van Hout BA. Discounting costs and effects: A reconsideration. Health Econ. 1998;7:581-594.

[35] Fenwick E, Claxton K, Sculpher MJ. The value of implementation and the value of information: Combined and uneven development. Med Decis Making. 2008;28:21-32.

[36] Barton GR, Briggs AH, Fenwick EAL. Optimal cost-effectiveness decisions: The rule of the cost-effectiveness acceptability curve (CEAC), the cost-effectiveness acceptability frontier (CEAF), and the expected value of perfection information (EVPI). Value in Health. 2008;11:886-897.

[37] Fenwick E, Claxton K, Sculpher MJ. Representing uncertainty: The role of cost-effectiveness acceptability curves. Health Econ. 2001;10:779-787.

[38] Fenwick E, O'Brien BJ, Briggs A. Cost-effectiveness acceptability curves – facts, fallacies and frequently asked questions. Health Econ. 2004;13:405-415.

[39] Strong M, Oakley JE, Brennan A. Estimating multiparameter partial expected value of perfect information from a probabilistic sensitivity analysis sample: A nonparametric regression approach. Med Decis Making. 2014;34:311-326.

[40] Cohen MA, Piquero AR. New evidence on the monetary value of saving a high risk youth. Journal of Quantitative Criminology. 2009;25:25-49.

[41] Cohen MA, Piquero AR, Jennings WG. Studying the costs of crime across offender trajectories. Criminology & Public Policy. 2010;9:279-305.

[42] OECD Library (2015) Purchasing power parities for GDP. Available: http://www.oecd-ilibrary.org/ economics/purchasing-power-parities-for-gdp-2014-5_ppp-gdp-table-2014-5-en;jsessionid = j1wryvqunpvk.x-oecd-live-03. Accessed 1 March 2015.

[43] McKenna C, Claxton K. Addressing adoption and research design decisions simultaneously: The role of value of sample information analysis. Med Decis Making. 2011;31:853-865.

[44] Claxton K, Palmer S, Longworth L, Bojke L, Griffin S, et al. Informing a decision framework for when NICE should recommend the use of health technologies only in the context of an appropriately designed programme of evidence development. Health Technol Assess. 2012;16.

[45] Netherlands Youth Institute. Database of effective youth interventions. Available: http://www.nji. nl/nl/Kennis/Databanken/Databank-Effectieve-Jeugdinterventies/ Erkende-interventies. Accessed 31 January 2014.

[46] Walker S, Sculpher M, Claxton K, Palmer S. Coverage with evidence development, only in research, risk sharing, or patient access scheme? A framework for coverage decisions. . Value in Health. 2012;15:570-579.

[47] Bojke L, Claxton K, Sculpher M, Palmer S. Characterizing structural uncertainty in decision analytic models: A review and application of methods. Value in Health. 2009;12:739-749.

[48] Briggs AH. Handling uncertainty in cost-effectiveness models. Pharmacoeconomics. 2000;17:479-500.

[49] Bilcke J, Beutels P, Brisson M, Jit M. Accounting for methodological, structural, and parameter uncertainty in decision-analytic models: A practical guide. Med Decis Making. 2011;31:675-692.

[50] Haij Ali Afzali H, Karnon J. Exploring structural uncertainty in model-based economic evaluations. Pharmacoeconomics. 2015;20 January 2015.

[51] Frederix GWJ, Van Hasselt JGC, Schellens JHM, Hövels AM, Raaijmakers JAM, et al. The impact of structural uncertainty on cost-effectiveness models for adjuvant endocrine breast cancer treatments: The need for disease-specific model standardization and improved guidance. Pharmacoeconomics. 2014;32:47-61.

[52] Chassin L, Piquero AR, Losoya SH, Mansion AD, Schubert CA. Joint consideration of distal and proximal predictors of premature mortality among serious juvenile offenders. J Adolesc Health. 2013;52:689-696.

[53] Teplin LA, McClelland GM, Abram KM, Mileusnic D. Early violent death among delinquent youth: A prospective longitudinal study. Pediatrics. 2005;115:1586-1593.

[54] Jackson CH, Bojke L, Thompson SG, Claxton K, Sharples LD. A framework for addressing structural uncertainty in decision models. Med Decis Making. 2011;31:662-674.

[55] McCollister KE, French MT, Fang H. The cost of crime to society: New crime-specific estimates for policy and program evaluation. Drug Alcohol Depend. 2010;108:98-109.

[56] Van Zutphen F, Goderie M, Janssen J. De maatregel Inrichting Stelselmatige Daders (ISD). Maatschappelijke kosten-batenanalyse van een eventuele verlenging [in Dutch]. Rotterdam, The Netherlands: Van Zutphen Economisch Advies; 2014.

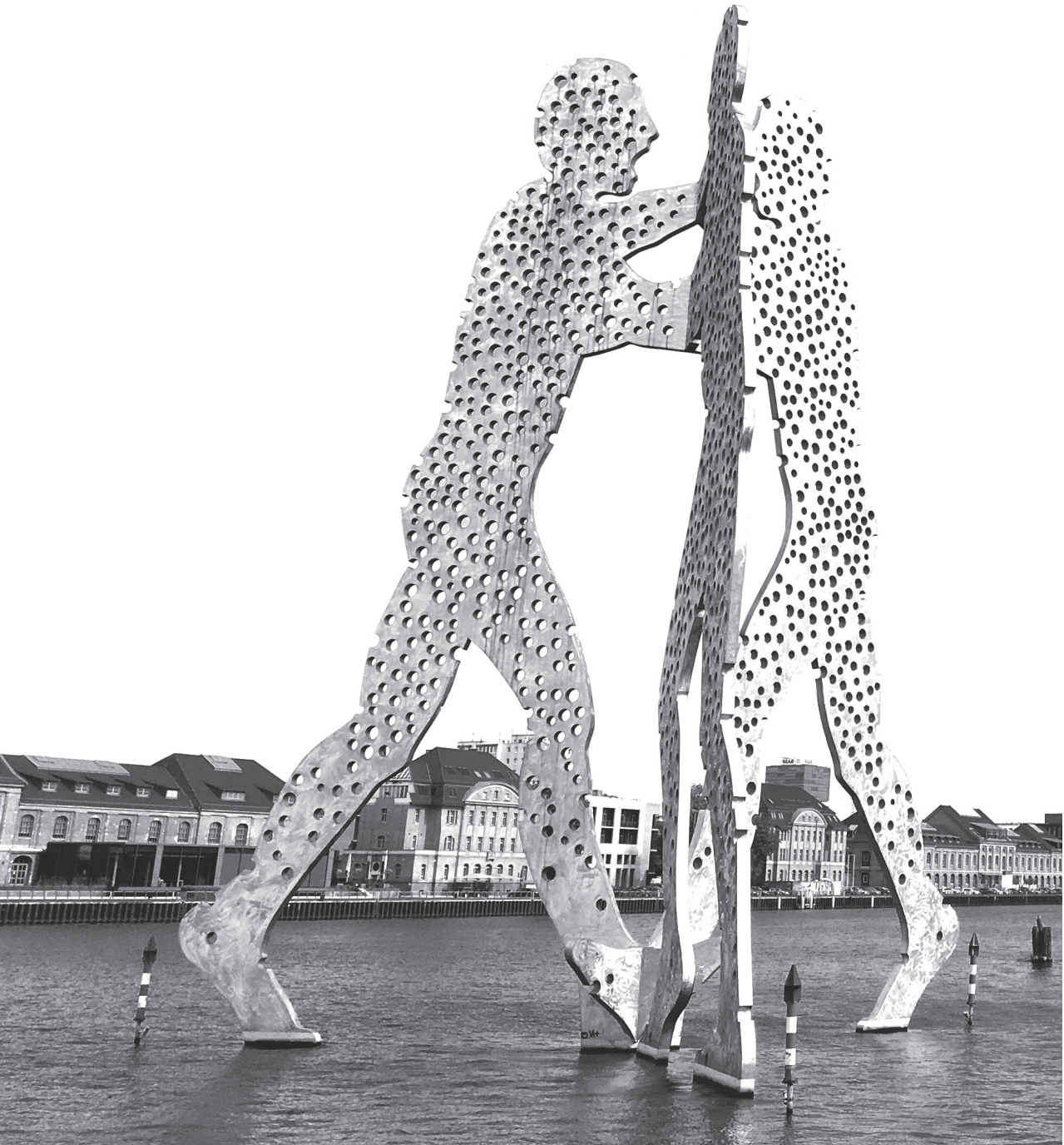
[57] Cohen MA, Rust RT, Steen S, Tidd ST. Willingness-to-pay for crime control programs. 2004;42:89-109.

[58] Moolenaar DEG. Modelling Criminal Justice System Costs by Offence. European Journal on Criminal Policy and Research. 2009;15:309-326.

[59] Aos S, Lieb R, Mayfield J, Miller M, Pennucci A. Benefits and costs of prevention and early intervention programs for youth. Olympia: Washington State Institute for Public Policy; 2004.

[60] Versantvoort MC, Verster ACM, Jannink J, Van den Broek LGJM, Van Zutphen F, et al. Kosten en baten van justitiele interventies. Ontwikkeling van een analyse- en rekenmodel [in Dutch]. Rotterdam, The Netherlands: ECORYS; 2005.

[61] Donker A, De Bakker W. Vrij na een PIJ. Voorspellende factoren van acceptatie vrijwillige nazorg en recidive na een PIJ-maatregel [in Dutch]. Leiden, The Netherlands: WODC, Ministerie van Veiligheid en Justitie, Hogeschool Leiden; 2012.



Chapter 6

Clinicians' Views on Therapeutic Outcomes of Systemic Interventions and on the Ability of the EQ-5D to Capture these Outcomes

Based on Schawo, S., Brouwer, W., Hakkaart-van Roijen, L.

The Journal of Mental Health Policy and Economics. 2017; 20(3): 131-136.

Background

In the light of budget constraints on mental health care expenditures, economic evaluations have increasingly gained importance in mental health care. Yet, it is unclear whether existing methodology used in health economic evaluations captures effects of interventions in mental health equally well as those of purely medical interventions. This particularly holds for interventions, which focus on broader improvements rather than measurable changes in health alone [1, 2]. The goals of these interventions may be broad, individualized and may not easily be assessed in dimensions of medical functioning such as those of a common outcome measure like the EuroQol five dimensions questionnaire (EQ-5D) [2]. When compared to medical interventions, effects of mental health interventions may be defined less clearly in medical terms but rather as more broad improvements of well-being. Shah and colleagues indicate, based on a survey among the general UK public, that mental health may not be sufficiently captured by the EQ-5D [3].

Systemic interventions, as a particular example of mental health interventions, focus on broad improvements of the client's functioning within his environment. These mental health interventions are increasingly used to treat mental disorders in adolescents. Multidimensional Family Therapy (MDFT), Functional Family Therapy (FFT), Multisystemic Therapy (MST) are examples of such interventions. All systemic interventions have in common that they are particularly directed at improving the interaction between the client and his environments [4] and hence involve not only medical improvements of the client but also improvements in communication and interaction with family and often peers and school in the treatment process [5, 6]. Measurement of these effects may be relevant to the evaluation of treatment success. There are some differences as to the target population of the different types of systemic interventions. MDFT is in particular used to treat adolescents with substance use disorders and related problems [7]. MST is mainly used to treat violent behavior [6]. FFT is predominantly used to treat slightly less serious cases of juvenile delinquency [5] as it is a less intensive program when compared to MST. Systemic therapy in general is also used in internalizing disorders, for cases where involvement of the system around the client is of importance. As all systemic interventions involve the close environment of the client, effects on (interaction with) third parties can reasonably be expected. Also the nature of the treated disorders itself (i.e. substance use, delinquency) may imply broad effects and often involves

comorbidities [8], which make the effects of these disorders less clear-cut than those of other, purely medical, disorders.

In health economic evaluations, the costs of interventions are set off against their effects. According to the health economic guidelines [9, 10], these effects are commonly measured in terms of quality adjusted life years (QALY). QALYs capture effects in terms of health-related quality and quantity of life. Hence QALYs are particularly suitable to assess medical effects. QALYs are derived with generic instruments such as the EQ-5D [11] or the Short-Form 6-Dimension[12], which are completed by patients. For these instruments societal preference-weights are available, hence it is possible to weigh different aspects of patient responses about their health status based on the societal judgment of the importance of these aspects and convert these into one single number (i.e. QALYs). Of the available instruments, the EQ-5D [11] is the preferred instrument in Europe as it has been translated in various languages, has a wide range of local societal value sets available and provides a single index of health-related quality of life [13, 14]. It contains a descriptive part and a visual analogue scale (VAS) and is particularly useful in measuring health-related treatment effects. The descriptive part, on which this paper focuses, includes five dimensions (i.e. mobility, pain/discomfort, self-care, usual activities, depression/anxiety) with three levels each (i.e. no, some or extreme problems) [11].

Studies on the suitability of the use of the EQ-5D in several mental health populations provide both positive and negative evidence for its use. In populations of schizophrenia patients, Willige [15] stated that the EQ-5D score did not sufficiently reflect changes in social and psychological well-being. Yet, Prieto [16] suggested that the EQ-5D was valid to be used to assess different degrees of illness in schizophrenia patients. In a population of chronic heroine-dependent-patients van der Zanden [17] found the EQ-5D a suitable measure. In bipolar disorder, Hayhurst [18] considered the EQ-5D useful in measuring symptoms of depressions, however its ability to reflect manic symptoms could not be shown due to a limited sample of investigated patients. Pyne [19] presented evidence for the use of health-related quality of life instruments in populations of substance use disorders but noted that problems with legal issues and alcohol were not properly reflected by the measures. Coast [20] expressed more general criticism on the use of the QALY when assessing effects broader than health. In 2007, Knapp [2] expressed criticism on the ability of measures like the EQ-5D to properly assess changes in mental health conditions. He suggested that new

measures were needed as to assess interventions in mental health [2] focusing on the specific aspects of quality of life of different mental health problems. He expressed the necessity to use an instrument, which properly reflects the relevant improvements within the treatment process [2].

Hence, though the EQ-5D is commonly used in economic evaluations it is yet unclear if it sufficiently captures all relevant benefits of interventions in mental health care. To improve this situation, several options can be distinguished, including the development of a new outcome measure, the adjustment of an existing measure to allow inclusion in economic evaluation or the extension of the EQ-5D measure with 'bolt on' dimensions [3]. The latter entails adding additional dimensions to the current instrument in an effort to improve its evaluative scope and sensitivity. For all options, it is necessary to obtain information on the relevant domains to be measured (additionally). Hence, this study is an explorative study with the goal of gaining initial information on relevant domains for potential development or improvement of instruments, which would properly capture the effects of mental health interventions in the future and can be used in economic evaluations. As a first step to investigate this issue, this paper focuses on systemic interventions, which typically aim to achieve goals beyond health gains alone. This is done based on the views of clinicians on relevant domains and on the ability of the descriptive part of the EQ-5D to capture these domains. We considered clinicians a reliable source to describe the relevant and specific aspects of treatment success and we aimed to find out whether, from their perspective rather than from the perspective of health economists, the current health economic method of measuring effects based on the EQ-5D captures the most relevant treatment effects. Based on semi-structured interviews with the clinicians we explored the effects of systemic interventions. Then we inquired whether the clinicians considered the dimensions of the EQ-5D to fit and capture the relevant therapeutic goals and effects, and whether according to their view, the instrument missed dimensions in order to properly capture the goals of systemic interventions.

Methods

We used a qualitative research design of semi-structured interviews as common within the context of explorative analyses. We performed these semi-structured interviews to attain a first impression of the domains clinicians consider relevant in the evaluation of systemic interventions. Individual semi-structured interviews were conducted with seven clinicians at the mental health institutions

'Brijder' and 'de Viersprong' in the Netherlands. The interviews were held between the 24th of November 2011 and the 2nd of February 2012 at The Hague, Halsteren and Etten Leur.

The aim was to at least include one clinician with knowledge of a particular type of systemic interventions (MST, MDFT, FFT and general systemic family therapy) in order to have a complete picture of relevant outcomes across these types of systemic interventions. All clinicians were active providers of one of the systemic interventions to young patients from ages 12-18. Clinicians were approached via the team leader of their systemic intervention unit, who was provided with a general introduction to the research. Individual clinicians were asked to take part in an interview of approximately one hour. All interviews were fully recorded.

The individual interviews were started with a general introduction on the research project. Participants were informed that the interviews were to be recorded and that no reward would be given. The interviews included three parts of guided open questions. The first part was aimed at retrieving general information about the clinicians. The second part was intended to attain a general idea of the therapeutic goals within a systemic therapy setting.. In the third part of the interviews, the EQ-5D was introduced. As we inquired on the suitability of the instrument to capture relevant treatment effects this part of the interview was most relevant and central to our research question. The structure of the three parts of the interview is provided in table 1 below.

Table 1 | Structure of the semi-structured interviews with clinicians

Part 1: General characteristics of the clinician including age, gender, educational background, geographical region of work, type of systemic therapy that was provided, years of experience and the approximate number of clients seen per week;
Part 2: Most important outcomes of systemic interventions as perceived by clinicians;
Part 3: Clinicians' judgment on the ability of the EQ-5D questionnaire to capture relevant therapeutic goals of systemic interventions and the effects which clinicians may miss when the EQ-5D was used to evaluate the effects of systemic interventions. - Per EQ-5D dimension judgment on suitability to pick up effects of systemic interventions (plus necessary specifications per dimension) - Mention of possible missing dimensions when (solely) using the EQ-5D to measure outcomes of systemic interventions

We posed explorative questions and used inductive coding [21], performed by one researcher, to retrieve overarching categories or domains. Terms mentioned

by the clinicians were translated from Dutch to English and grouped in major categories.

Results

Seven clinicians were interviewed. The proportion of female respondents in the group of interviewed clinicians was relatively high with 71.4%, which is in line with the higher number of females within the profession of clinicians. The age of the clinicians ranged from 30 to 60 years with a mean of 42.1 years. Four out of the seven clinicians worked for the mental health institution Brijder, which is specialized in treatment of substance use related problems. The remaining three clinicians worked for the mental health institution 'de Viersprong', which is specialized in personality disorders. Geographical regions in which the clinicians worked included four Dutch provinces: North Holland, South Holland, Brabant and Zeeland. Six of the clinicians had a Bachelor's degree in Psychology and one had a Master's degree in psychology. Six had obtained training on systemic interventions and one was currently following training on systemic interventions to learn to perform a specific type of systemic intervention or systemic treatment in general (i.e. MDFT, FFT, MST, and general systemic therapy). The years of experience with systemic therapy of the clinicians ranged from 0.1 to 23.0 years with a mean of 8.4 years, and a median of 4.5 years. The number of clients seen per week ranged from 2.5 to 9.5 per week with a mean of 5.4 per week.

The responses to the question of what would be the most important outcomes of systemic interventions in practice were listed. Based on inductive coding answers of respondents were summarized in ten major categories. Items of the category 'family interaction/functioning' were most often mentioned with 38 sub-terms belonging to that category. Furthermore, 'parental functioning', 'criminal behavior of the youth', 'parental (mental) health', '(mental) health and functioning of the youth', 'substance use of the youth', 'social competences of the youth', 'school or work attendance of the youth' and 'marital functioning of parents' were mentioned between four and eight times. Finally, 'financial problems of the parents' were mentioned twice.

The answers to the questions on the relevance of the EQ-5D dimensions provided clinicians' judgments on the ability of the current dimensions to capture relevant outcomes of systemic interventions and suggestions for additions per dimension. The interviewed clinicians considered several of the EQ-5D dimensions relevant for the evaluation of systemic interventions. The dimensions 'usual activities'

and 'anxiety/depression' were considered particularly important to measure the outcome of systemic interventions. Respectively seven and six respondents considered these relevant. The dimension 'self-care' was judged relevant by three out of the seven respondents. The dimensions 'pain/discomfort' and 'mobility' were considered less relevant. Five clinicians considered 'pain/discomfort' irrelevant. All clinicians judged the dimension 'mobility' to be irrelevant as well. The clinicians further suggested specific textual additions to the existing EQ-5D dimensions, which could make the instrument more suitable for assessment of systemic interventions. Clinicians suggested the dimension 'usual activities' to focus on school, leisure, and work and on meeting appointments or agreements. According to the clinicians, the dimension 'anxiety/depression' should explicitly include substance-related aspects of depression and anxiety and aspects of aggression. The 'self-care' dimensions should specify and include not only daily aspects such as being able to get up, brushing teeth and taking regular meals but also more general aspects such as maintaining a healthy pattern, adhering to medication, being able to use a bike or public transport. Dimensions which the clinicians considered relevant for the evaluation of systemic interventions and which are not captured by the EQ-5D instrument were aspects of family functioning, systemic relations (peer, school, other), addiction, parental functioning and mental health, daily functioning (useful activities and occupation), aggression and self-confidence. According to the interviewed clinicians these aspects were not sufficiently covered by the EQ-5D.

Discussion & Conclusions

Based on semi-structured interviews, this study aimed to explore clinicians' views on the therapeutic goals of systemic interventions and to elicit their opinion on the ability of the EQ-5D to sufficiently capture and evaluate these goals. This study has provided a first indication of the aspects, which, from the view of clinicians, should be considered when evaluating the cost-effectiveness of systemic interventions.

The interviews revealed that clinicians considered a broad array of outcomes relevant to the measurement of the effect of systemic interventions. These outcomes not only included medical aspects, but also encompassed broader (societal) effects such as family functioning, parental functioning, social competencies, school attendance, etc. The clinicians considered several of the EQ-5D dimensions relevant (i.e. in particular 'usual activities' and 'anxiety/depression') for the evaluation of systemic interventions. Nonetheless, they

suggested that some of the dimensions needed additional specifications in order to be more suitable in this context. This does imply moving away from the current content of those dimensions, as included in the EQ-5D. Furthermore, without an exception all interviewed clinicians emphasized that a number of broader life dimensions were missing in the EQ-5D but relevant (and required) when evaluating systemic interventions. Most often mentioned by the clinicians in this context were systemic dimensions such as family relations and relations with others (peer, school, etc). Also specific aspects of addiction were considered relevant additions by several of the clinicians.

Given these findings, one could choose different paths to attain a suitable instrument for evaluation of systemic interventions. One of these could be the use of bolt-ons to the EQ-5D or mapping specific, non-preference based instruments on the EQ-5D. Also one could search among existing instruments for an instrument more sensitive to the underlying goals of the interventions. Such an instrument would, besides the common medical aspects, also capture the systemic aspects of the interventions and aspects of addiction. There are several instruments available in the field of addiction and delinquency which assess multiple dimensions of functioning of a client within his environment and which may hence be suitable to be used. Examples of such instruments may be the Adolescent Drug Abuse Diagnosis (ADAD) [22], the Child Adolescent Functional Assessment Scale (CAFAS) [23], the Global Appraisal of Individual Needs (GAIN) [24], the Teen Addiction Severity Index (T-ASI) [25], and the WAJCA-RA structured interview [26]. However, none of these instruments was designed for use in economic evaluations. Hence, no preference weights (utilities) for their outcomes exist, severely limiting their use in economic evaluations. A possible route forward would be to derive such preference weights for more elaborate instruments, which would facilitate their use in the context of economic evaluations of systemic interventions (in a similar way as the EQ-5D, i.e., in cost-utility analyses). This does imply that their results are not comparable to common economic evaluations using a different concept of relevant outcome (health-related quality of life) and a different outcome measure (e.g., EQ-5D). An alternative would be to search for broader outcome measures that could be relevant in both contexts, such as wellbeing measures.

Whichever way one chooses to go forward, our overall conclusion from this study is that more dimensions should be included in the evaluation of systemic interventions than is currently done.

The current study was an explorative analysis and has clear limitations affecting the generalizability of the presented results. First, the total number of therapists was small as we had to deal with practical issues of inclusion (such as limited time and availability of therapists). The selection of respondents therefore was partly pragmatic. However, we included at least one therapist per type of systemic intervention to cover all types of systemic interventions provided by the institutions. Furthermore, the number of interviewed therapists included a larger proportion of clinicians with specialization on MDFT. Hence there was a(n intentional) strong focus on clients with problems of substance use. A consequence may be that the importance of addiction-related improvements may have been overemphasized. Yet, we observed that, despite the small sample size, the dimensions we found in this study are in line with the literature on the goals of systemic interventions [6; 27; 28; 29]. We also observed considerable overlap between the answers of the different therapists, suggesting saturation so that additional interviews would not result in additional dimensions or insights. Furthermore, one of the interviewed clinicians had just started providing systemic interventions and, therefore, had little experience with these interventions. In addition, the current study was set up as a pilot study aiming to perform an explorative analysis of the goals of systemic interventions and of the ability of the EQ-5D to capture these effects in economic evaluations. Hence, the results need to be interpreted with this (explorative) intention in mind.

Practical implications of the current study are that enhancements of the current health economic methodology appear necessary when evaluating systemic interventions. To capture all relevant outcomes influenced by these interventions in economic evaluations, in particular broader outcome measures than purely health-related quality of life measures such as the EQ-5D seem required. A focus of future research could be on investigating the suitability of other available instruments for use in economic evaluations of systemic interventions or to make existing (validated) instruments like T-ASI or ADAD suitable for this purpose. Without appropriate outcome measures, evaluations may risk misinforming policy makers and funding decisions.

References

[1] Ryan M, Gerard K. Using discrete choice experiments to value health care programmes: current practice and future research reflections. Appl Health Econ Health Policy 2003; 2(1): 55-64.

[2] Knapp M, Mangalore R. The trouble with QALYs... . Epidemiol Psychiatr Soc 2007; 16(4): 289-293.

[3] Shah KK, Mulhern B, Longworth L, Janssen MF. Views of the UK General Public on Important Aspects of Health Not Captured by EQ-5D. Patient 2017; 13:1-9.

[4] Brysbaert M. Psychologie Gent: Academia Press, 2009.

[5] Cottrell D, Boston P. Practitioner review: The effectiveness of systemic family therapy for children and adolescents. J Child Psychol Psychiatry 2002; 43(5): 573-586.

[6] Randall J, Cunningham PB. Multisystemic therapy: a treatment for violent substance-abusing and substance-dependent juvenile offenders. Addict Behav 2003; 28(9): 1731-1739.

[7] Liddle HA, Dakof GA, Parker K, Diamond GS, Barrett K, Tejeda M. Multidimensional family therapy for adolescent drug abuse: results of a randomized clinical trial. Am J Drug Alcohol Abuse 2001; 27(4): 651-688.

[8] Wicks-Nelson R, Israel AC. Abnormal Child and Adolescent Psychology New York: Pearson, 2013.

[9] The Health Care Insurance Board (CVZ). Guidelines for Pharmacoeconomic Research [in Dutch] Diemen: CVZ, 2006.

[10] Drummond MF, Sculpher MJ, Torrance GW, O'Brien BJ, Stoddart GL. Methods for the economic evaluation of health care programmes Oxford University Press, 2005.

[11] Brooks R, Rabin R, De Charro F. The measurement and valuation of health status using EQ-5D: a European perspective: evidence from the EuroQol BIOMED Research Programme Springer, 2003.

[12] Brazier J, Roberts J, Deverill M. The estimation of a preference-based measure of health from the SF-36. Journal of health economics. 2002; Mar 31;21(2):271-92.

[13] Rabin R, Charro FD. EQ-SD: a measure of health status from the EuroQol Group. Annals of medicine. 2001; Jan 1;33(5):337-43.

[14] Szende AG. EQ-5D value sets: inventory, comparative review and user guide. Oppe M, Devlin NJ, editors. Dordrecht: Springer; 2007.

[15] van de Willige G, Wiersma D, Nienhuis FJ, Jenner JA. Changes in quality of life in chronic psychiatric patients: a comparison between EuroQol (EQ-5D) and WHOQoL. Qual Life Res 2005; 14(2): 441-451.

[16] Prieto JM, Atala J, Blanch J, Carreras E, Rovira M, Cirera E, et al. Psychometric study of quality of life instruments used during hospitalization for stem cell transplantation. J Psychosom Res 2004; 57(2): 201-211.

[17] van der Zanden BP, Dijkgraaf MG, Blanken P, de Borgie CA, van Ree JM, van den Brink W. Validity of the EQ-5D as a generic health outcome instrument in a heroin-dependent population. Drug Alcohol Depend 2006; 82(2): 111-118.

[18] Hayhurst H, Palmer S, Abbott R, Johnson T, Scott J. Measuring health-related quality of life in bipolar disorder: relationship of the EuroQol (EQ-5D) to condition-specific measures. Qual Life Res 2006; 15(7): 1271-1280.

[19] Pyne JM, French M, McCollister K, Tripathi S, Rapp R, Booth B. Preference-weighted health-related quality of life measures and substance use disorder severity. Addiction 2008; 103(8): 1320-9; discussion 1330-2.

[20] Coast J, Flynn TN, Natarajan L, Sproston K, Lewis J, Louviere JJ, et al. Valuing the ICECAP capability index for older people. Soc Sci Med 2008; 67(5): 874-882.

[21] Kvale S, Brinkmann S. Interviews: An Introduction to Qualitative Research Interviewing. London: Sage Publications, 2009.

[22] Friedman, AS, Utada A. A method for diagnosing and planning the treatment of adolescent drug abusers (the Adolescent Drug Abuse Diagnosis [ADAD] instrument). J Drug Educ 1989; 19(4): 285-312.

[23] Hodges K, Wong MM. Psychometric Characteristics of a Multidimensional Measure to Assess Impairment: The Child and Adolescent Functional Assessment Scale. J Child Fam Stud 1996; 5(4): 445-467.

[24] Dennis ML. Global Appraisal of Individual Needs (GAIN): Administration guide for the GAIN and related measures (Version 1299) Bloomington, IL: Chestnut Health System, 1999.

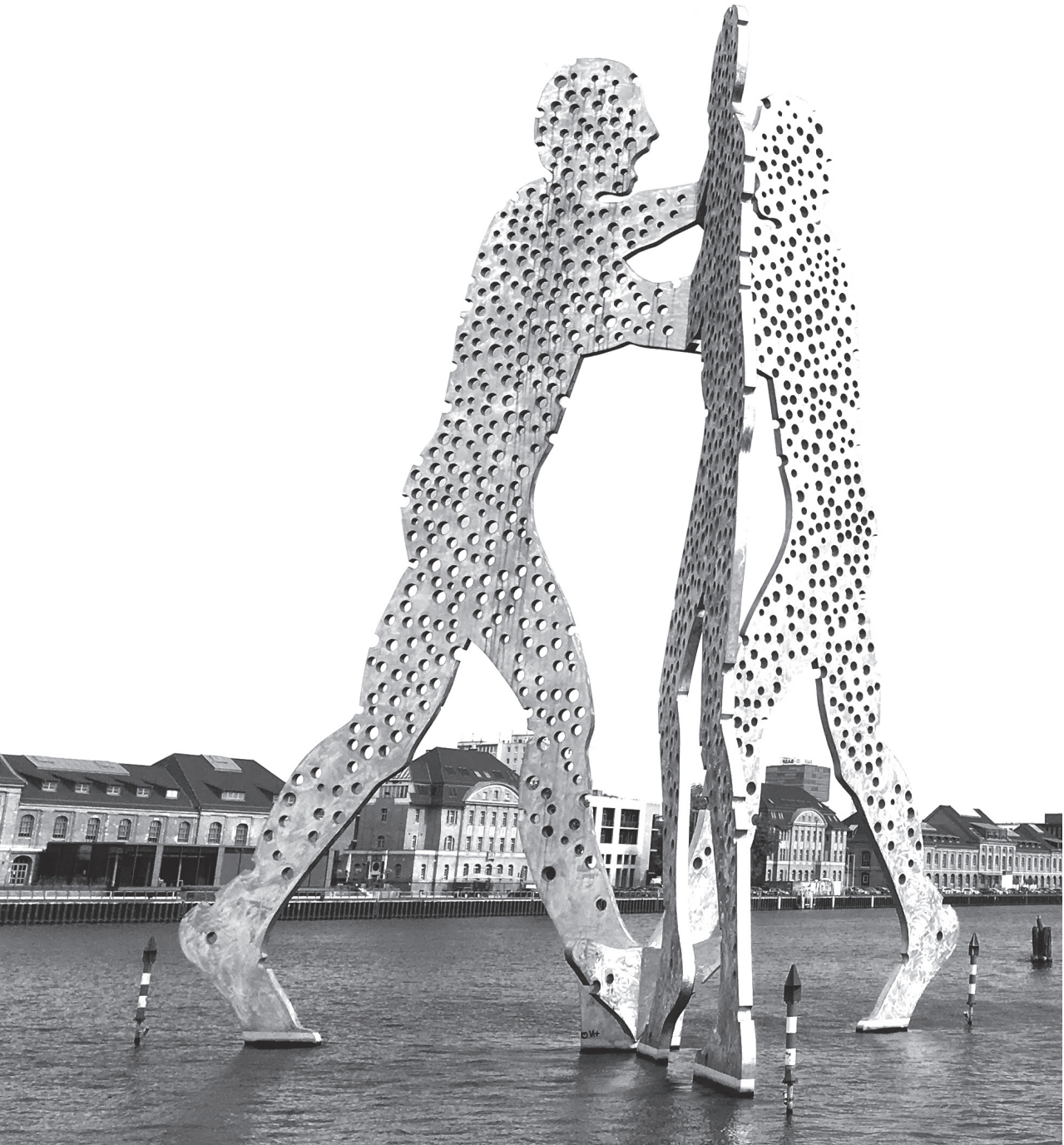
[25] Kaminer Y, Bukstein O, Tarter RE. The Teen-Addiction Severity Index: rationale and reliability. Int J Addict 1991; 26(2): 219-226.

[26] Washington State Institute for Public Policy. (1999). Washington State Juvenile Court Assessment Manual Version 2.0. 1999. http://www.wsipp.wa.gov/rptfiles/99-01-0000.pdf. Accessed 1 Oct 2013.

[27] Hendriks V, van der Schee E, Blanken P: Treatment of adolescents with a cannabis use disorder: Main findings of a randomized controlled trial comparing multidimensional family therapy and cognitive behavioral therapy in The Netherlands. *Drug Alcohol Depend* 2011, 119(1):64-71.

[28] Henggeler S W, Sheidow A J: Empirically supported family-based treatments for conduct disorder and delinquency in adolescents. *J Marital Fam Ther* 2002, 38(1): 30-58.

[29] French M T, Zavala S K, McCollister K E, Waldron H B, Turner C W, Ozechowski T J: Cost-effectiveness analysis of four interventions for adolescents with a substance use disorder. *J Subst Abuse Treat* 2008, 34(3): 272-281.



Chapter 7

The Search for Relevant Outcome Measures for Cost-utility Analysis of Systemic Interventions in Adolescents with Substance Use Disorder and Delinquent Behavior: A Systematic Literature Review

Based on Schawo, S., Bouwmans, C., van der Schee, E., Hendriiks, V., Brouwer, W., Hakkaart-van Roijen, L.

Health and Quality of Life Outcomes. 2017; 15(1):179.

Background

Systemic family interventions are psychotherapeutic treatments, which are increasingly used to treat children and adolescents with mental disorders. These interventions are based on the idea that the behavior of a patient is the result of interactions between himself and the different 'systems' he is involved in (i.e. family, peers, school, etc.) and of the interactions between these systems [1-3]. Treatment is directed at improving the disturbing aspects within these interactions [3] and it actively involves the systemic context of the patient. Hence, potential effects are broad and may range from improvements in the interactions with parents, other family members, peers or neighbors, to improvements in educational achievements and work relations, reduction of criminal activity and substance use and reduction of problems with the juvenile justice system [2, 4-6]. Systemic family interventions have shown particularly effective in the treatment of adolescents with substance use disorders and delinquency [7-10]. Examples of these interventions are Multisystemic Therapy (MST), Functional Family Therapy (FFT), Multidimensional Family Therapy (MDFT) and Brief Strategic Family Therapy (BSFT) [7-10].

With the increasing use of systemic family interventions, the question of funding and reimbursement arises. In some countries, like the Netherlands or the United Kingdom, systemic family interventions are reimbursed from social health insurance schemes and, as such, are part of collectively financed health care. Hence, the interventions compete for limited funds with other health care expenditures and, on top of proving effective, need to demonstrate value for money. Common practice in the economic evaluation of medical interventions is the use of cost-utility analysis (CUA) [11, 12] measuring effects in terms of Quality-Adjusted Life-Years (QALYs). QALYs combine length and quality of life gained. Typically, quality of life is measured through preference-based, generic health outcome measures (such as the EQ-5D). These outcome measures typically concentrate on improvements in a number of health domains. A recent publication of our department [13] described the results of a CUA of MDFT versus Cognitive Behavioral Therapy (CBT) in which the effects were measured with the EQ-5D. Yet, in the field of mental health, doubts have been expressed [14, 15] on the use of these generic quality of life measures [16] as these tools might be too limited to cover all relevant treatment effects. Studies on the applicability of these measures in mental health have presented mixed results [14, 15]. Furthermore, there is increasing attention for the inclusion of spillover effects on caregivers and families in economic evaluations. Currently, these effects are

not yet included [17, 18], though they may be particularly important in treatment of younger patients. Recently, the Second Panel of Cost-Effectiveness in Health and Medicine has recommended further research on quality of life effects on family members of patients [19].

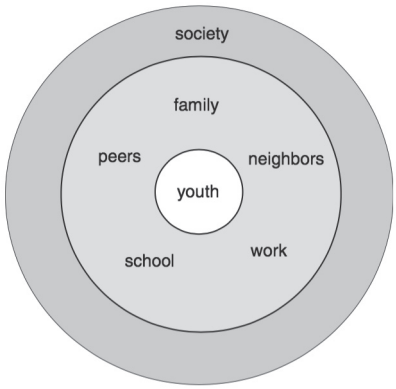
Both aspects, the assessment of effects specific to mental health treatments and the inclusion of (partial) effects on third parties, seem of particular relevance to the economic evaluation of systemic interventions in delinquency and substance use in adolescents. As outcomes of systemic family interventions are broad and transcend health gains, conventional CUA outcome measures may be too limited and insufficiently connected to clinical practice. This may be one of the reasons why economic evaluations of systemic interventions are still scarce and overall of low quality [20]. Existing economic evaluations of these interventions vary in setting, design and in outcomes measured [20] hence limiting the comparability of results. Furthermore, few studies consider effects on others than the patient [1].

If the aim is to perform economic evaluations of systemic family interventions which account for all relevant effects, a disorder-specific multidimensional measure that captures all relevant systemic contexts would be desirable. Ideally, if such a measure had societal preference-weights attached to its dimensions and levels, it would deviate from the common CUA methodology yet enable CUA-like economic evaluations. In patients with substance use disorder (one of the patient groups treated with systemic family interventions), the need for such a single comprehensive outcome measure capturing the full benefits of treatments has been recognized before [21]. Deas and Thomas [22] and Hogue and Liddle [23] emphasized the necessity of assessing various outcomes beyond effects in the adolescent. In an illustrative pilot study, Jofre-Bonet and Sindelar [21] presented a first example of a preference-based measure for adult populations with substance abuse. However, that measure was not based on standard preference-elicitation techniques but the authors attached patient preference-weights to the eight main domains of the Addiction Severity Index (ASI) [24] by constructing a weight index.

In the current study, we take this line of research further by searching for a multidimensional outcome measure to evaluate systemic family interventions in the populations of adolescents with substance abuse disorder or problems of delinquency. Such a measure could facilitate CUAs of systemic family interventions and could either be based on existing effectiveness measures in

this field or fully designed anew. In both cases, the use of an existing measure or the design of a new measure, relevant domains would need to be identified. Based on consultation of the literature on systemic family interventions [1, 25, 26] the domains relating to aspects of the individual patient, family, school (or work) and other community environments (e.g. peers, neighbors) were considered most relevant to the evaluation of the interventions. Figure 1 provides a graphical illustration of these domains, which indicate where potential effects may occur. The strength of the impact on the different systems may obviously differ, depending on the exact underlying problems and other contextual factors.

Figure 1 | Systems involved in systemic family interventions for treatment of delinquency and substance-abuse in adolescents



We perform a systematic literature review to investigate and appraise available instruments in the field of adolescent delinquency and substance use, which cover the relevant domains and which are already accepted and validated in the field. We assess which of these instruments might be most suited to serve as a basis for a preference-based measure in CUA, based on characteristics like comprehensiveness, brevity, accessibility, psychometric properties, etc. Advantage of using an existing instrument would be its being established, accepted and validated in the field and known by clinicians. It would then only be necessary to add preference-weights to the domains to account for differences in impact of each domain. In this way we aim to contribute to the development of adequate outcome measures to assess the economic value of systemic family interventions in the treatment of delinquency and substance use.

Methods

We conducted a systematic literature review to identify instruments within the effectiveness and efficacy literature of mental health interventions for adolescents with substance use disorder and delinquency problems. We then assessed the suitability of these instruments for use of preference elicitation techniques. The assessment was based on several characteristics relevant to attain societal preference weights. These characteristics were among others the coverage of the systems displayed in figure 1 (i.e. youth, family, peers, school, work, society and neighbors), brevity, practicability of use, accessibility, psychometric properties and acceptance in the field. The review protocol was not registered. Yet, this study adhered to the PRISMA reporting guidelines [27].

Criteria for inclusion

Types of participants

The target population of the systematic literature review consisted of adolescents between 12 and 18 years of age with symptoms of delinquency and/or substance use. Patients from specific sub-groups (e.g. homeless or runaway adolescents or adolescents with substance use disorder and comorbid depression) were excluded. As studies focusing on these subgroups evaluated specific outcomes, which were not necessarily relevant for the entire population of adolescents with substance use disorders and delinquent behavior, these studies were not considered relevant for the current study.

Types of interventions

We included studies on various mental health interventions for adolescents with substance use disorder or delinquency in a therapy/counseling setting in the systematic search to cover as many instruments as possible in the relevant target population. Individual interventions as well as systemic family interventions were included. Examples of such interventions are Cognitive Behavioral Therapy (CBT), Motivational Enhancement Therapy (MET), Multidimensional Family Therapy (MDFT), Multi Systemic Therapy (MST), Functional Family Therapy (FFT) and Ecologically Based Family Therapy (EBFT). Two types of interventions were excluded. First, interventions in mental health care that consisted of only pharmacotherapy were excluded since the focus of our study was specifically on the effect of psychosocial interventions. Second, mental health interventions for the prevention of criminal behavior or substance use disorder were excluded, as the symptoms within this group (i.e. high risk behavior or general behavioral

problems) were not considered severe enough to fit the definition of the target population.

Types of outcome measures

Our objective was to identify a wide array of instruments used to measure the effect of mental health interventions for adolescents with substance use disorders and delinquent behavior. Hence, we included studies with all measures of effectiveness and treatment outcome as well as efficacy studies.

Search methods for identification of studies

Databases were selected as to cover both interventions in the medical and in the educational field. The systematic literature review was performed in PubMed, Psychnet (PsycBOOKSc, PsycCRITIQUES, print), Cochrane and ERIC (Education Resource Information Center) to identify all effectiveness studies of mental health interventions for adolescent with substance use disorder or problems of delinquency. The databases were consulted between 5 March 2013 and 8 March 2013. Additional studies were identified based on reference list search. There were no restrictions on the type of publication. The language of publication was required to be English and publication date was 1990 or more recent. The search strategy used is displayed below.

(assessment OR measurement OR measure OR instrument OR "quality of life" OR effectiveness OR "well-being" OR outcome OR response)

AND

(delinquency OR crime OR criminal OR "substance abuse" OR substance OR cannabis OR marijuana OR alcohol OR cocaine OR heroin)

AND

(juvenile OR adolescent OR youth)

AND

(therapy OR intervention OR "family therapy" OR "family intervention" OR "functional family therapy" OR "multisystemic therapy" OR "multidimensional family therapy" OR "family-based therapy" OR "family treatment program" OR "mental health intervention" OR "mental health therapy" OR "cognitive behavioral therapy" OR "cognitive behavioural therapy" OR "psychosocial treatment" OR "contingency management" OR "voucher-based" OR FFT OR MST OR MDFT OR CBT)

Data analysis

Study selection

First, duplicates were removed. Then, the study selection was performed in two rounds. First, a selection based on title and abstract was performed, then selected articles were subject to a second screening based on full texts. Both rounds

of selection were performed by two researchers independently and were each followed by a round of consensus. The eligibility criteria for the first selection based on title and abstract were the following.

No duplicates

Language: English
Year: 1990 or more recent
Intervention: mental health intervention[a]
Disease/symptoms: delinquency, substance abuse (drug, alcohol)[b]
Population: adolescent (12-18 years)
Principle outcomes: effectiveness, efficacy, treatment outcome, instrument

[a]excluding interventions consisting of pharmacotherapy only and prevention interventions
[b]excluding less severe symptoms like high-risk behavior or general behavioral problems and specific subgroups (i.e. homeless, runaway, patients with comorbid depressive disorder)

Subsequently, when abstracts or titles adhered to the above screening criteria, full texts were independently screened for inclusion based on the following (additional) criteria.

Availability of full text

Intervention: mental health intervention[c]
Disease/symptoms: delinquency, substance abuse (drug, alcohol)[d]
Principle outcome: effectiveness, efficacy, treatment outcome, instrument[e]

[c]excluding interventions consisting of pharmacotherapy only and prevention interventions
[d]excluding less severe symptoms like high-risk behavior or general behavioral problems and specific subgroups (i.e. homeless, runaway, patients with comorbid depressive disorder)
[e]clear outcome domain or instrument was stated in the text; process measures such as therapy dose, therapy adherence or motivation to change were not considered principle outcomes and were hence excluded.

Furthermore, articles from reference lists of reviews were identified. For these, we performed a shortened screening and selection procedure. Titles of these articles were screened based on the following criteria: a) >=1990; b) peer-reviewed article; c) randomized control trial or effect/effectiveness/efficacy study/treatment outcome; d) adolescents; e) delinquency/offenders/substance-abuse; f) mental health intervention (no pharmacotherapy). If this selection resulted in inclusion, the abstract was screened and a final decision on inclusion or exclusion was made. Included articles were added to the database of identified articles for further data synthesis.

Data extraction

Data extraction was performed in MS Access with predefined fields. From all selected studies, general information, such as the title of the study, the name of the author, journal, etc., were recorded, as well as information on the sample size, the studied population and type of intervention (systemic, other [i.e. individual, group intervention], both).

In addition to this general information, instrument-specific information was extracted. This information consisted of instrument names (e.g. Child Behavior Checklist [CBCL]) and covered domains (e.g. family functioning, adolescent behavior, etc.). This information was recorded in order to identify the instruments currently used in the field and their coverage of the different systems relevant for the evaluation of systemic family interventions (figure 1).

Synthesis and evaluation of results

As a next step, domain names of the instruments were extracted from the identified articles and linked to the systems relevant for the evaluation of systemic family interventions (figure 1): youth, family, peers, school, work, society and neighbors. Domain names were verified with available resources such as guidelines, websites of the developer and other articles using the same instrument. After verification, the domains were translated into the systems mentioned in figure 1. For this purpose, domains related to the adolescents themselves, such as 'substance use and abuse', 'physical health' or 'mental health' were linked to the system 'youth' whereas domains such as 'family relations' were recoded into the system 'family', domains like 'peer relations', 'social skills' or 'leisure/recreation' were labeled as 'peer' system, domains like 'educational status' were labeled 'school' and 'delinquency' as 'society'. Table 1 provides an example of the process of recoding for the Problem Oriented Screening Instrument for Teenagers (POSIT). Next, all instruments were classified based on the number of systems (presented in figure 1) covered and ranked from highest to lowest. Those covering five or more systems were considered most relevant for our purpose as those covered the majority of effects of systemic family interventions in adolescents with substance use disorder or problems of delinquency.

Table 1 | Example of recoding of domains into systems

Problem Oriented Screening Instrument for Teenagers (POSIT)	
domain	**corresponding system**
Substance use and abuse	youth
Physical health	youth
Mental health	youth
Family relations	family
Peer relations	peers
Educational status	school
Vocational status	work
Social skills	peers
Leisure/recreation	peers,
Aggressive behavior/delinquency	society

In line with our aim to identify an instrument, which captures most of the systems relevant to the evaluation of systemic family interventions, those instruments covering more than five systems were evaluated in more detail. These were then appraised according to necessarily arbitrary characteristics of brevity, feasibility, practicability, accessibility, psychometric properties and acceptance in the field. These characteristics were set up as to identify one or more instruments suitable to attain societal preference-weights for an instrument by means of preference-elicitation techniques. Within preference-elicitation techniques, such as discrete choice experiments, the number of domains rarely exceeds ten [28, 29]. With higher numbers of domains, the decision task may become too complex and cognitively demanding for the respondent [28]. Hence, a suitable instrument should possess less than 10 domains. A second consideration was the practical use of the instrument itself in clients. An instrument, ideally suitable for self-completion, should put as little strain as possible on the respondent, without loss of important content. Hence, we set a limit to the maximum number of items of the instrument at 500 and a maximum completion time of 1 hour, assuming that these would be reasonable amounts of items and time to ask from respondents. Another criterion was the accessibility of the instrument as to ascertain ease of use in future studies. Evaluation of this criterion included the price of use and availability of a (digital) version. Psychometric properties were considered to judge the suitability of the instrument for integration in health economic evaluations. Findings from existing publications on validity and reliability of the instruments were considered in this context. Finally, the frequency of use of the instrument was considered an indicator for the acceptance of the instrument

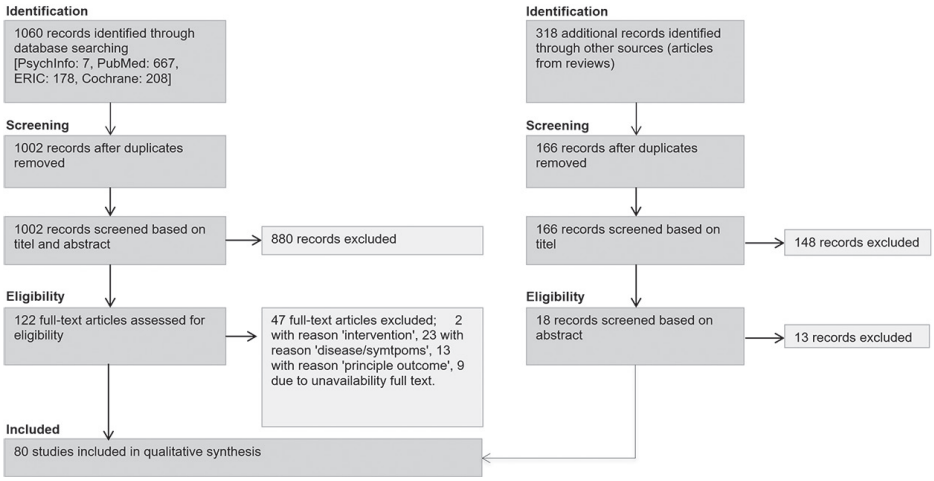
in the clinical field. This was approximated by the number of times that an instrument was used in the studies identified in this review.

Results

Study selection

The systematic search resulted in 1,060 articles. After duplicates were removed 1,002 articles remained. Screening based on abstracts resulted in the exclusion of 880 articles. Full text assessment of the remaining 122 articles resulted in the exclusion of 2 articles not matching the definition of the intervention, 23 articles not matching the disease or symptoms of the target population, 13 not matching the requirements for the principle outcome of the studies, and 9 due to unavailability of a full text version. Hence 75 articles were included. Furthermore, 318 underlying articles from reviews were screened. From these, 166 articles remained after duplications with the first search results were removed. The screening of these articles in a first round by title and in a second round by title and abstract resulted in the exclusion of 161 articles and inclusion of 5 additional publications (figure 2).

Figure 2 | Phases of the systematic review adapted from Moher et al. [28]



Study Results

A total of 80 articles were included in the synthesis. The aim was to identify clinical instruments in the field suitable for integration in a health-economic framework based on criteria of coverage of relevant systems, feasibility to perform preference-elicitation techniques, practicability of use, accessibility for future studies, psychometric properties and acceptance in the field. A

summary of the identified reviews and clinical trials is provided in tables 2 and 3 respectively. From the 80 selected articles we identified a total of 102 instruments, differing substantially in what these intended to measure and in whom. These instruments measured varying (combinations of) outcomes such as substance use, physical health, mental health, family relations, peer relations, school and work status and criminal history.

Table 2 | List of identified reviews

ID	Authors	Year	Population
1	Armelius Bengt-Åke, Andreassen Tore Henning	2007	youth with antisocial behavior
2	Baldwin SA, Christian S, Berkeljon A, Shadish WR.	2011	adolescent delinquents and substance-abusers
3	Borduin CM.	1999	criminal and violent adolescents
4	Brown SA, D'Amico EJ.	2003	adolescent substance abusers
5	Cottrell D, Boston P.	2002	patients with conduct and attention deficit disorders, substance misuse, etc.
6	Curtis N, Ronan K, Borduin C M	2004	antisocial youths and youths with serious emotional disturbances
7	Deas D, Thomas SE.	2001	adolescents with substance use disorders
8	Deas D.	2007	adolescents with AOD disorders
9	Diamond G, Josephson A.	2005	adolescent substance use
10	Ferguson LM, Wormith JS.	2012	(adult and) young offenders
11	Henggeler SW, Sheidow AJ.	2012	conduct disorder and delinquency in adolescents
12	Henggeler SW, Sheidow AJ.	2003	conduct disorder and delinquency in adolescents
13	Hogue A, Liddle HA.	2009	adolescent substance abuse
14	Littell Julia H, Campbell Margo, Green Stacy, Toews Barbara	2005	(among others) delinquent youth
15	Randall J, Cunningham PB.	2003	violent substance-abusing and substance-dependent juvenile offenders
16	Tanner-Smith EE, Wilson SJ, Lipsey MW.	2013	adolescent substance use disorder
17	Tripodi SJ, Bender K, Litschge C, Vaughn MG.	2010	adolescent alcohol use
18	Waldron HB, Kaminer Y.	2004	adolescent substance use disorders
19	Waldron HB, Turner CW.	2008	adolescent substance abuse
20	Walker D F, McGovern S K, Poey E L, Otis K E	2004	adolescent sexual offenders
21	Woolfenden Susan, Williams Katrina J, Peat Jennifer	2001	adolescents with delinquency or conduct disorder

Table 3 | Summary of identified clinical trials

ID	Authors	Year	Population	Age	N	Effect measures
22	Arnold EM, Kirk RS, Roberts AC, Griffith DP, Meadows K, Julian J.	2003	incarcerated, sexually abused adolescent females	12-17	100	Multidimensional Adolescent Assessment Scale (MAAS)
23	Borduin CM, Mann BJ, Cone LT, Henggeler SW, Fucci BR, Blaske DM, Williams RA	1995	juvenile offenders	12-17	176	Symptom Checklist-90 (SCL-90-R); Revised Behavior Problem Checklist (RBPC); Family adaptability and cohesion evaluation scales-II (FACES-II); Nine-item unrevealed differences questionnaire-revised; Peer relations inventory (MPRI)
24	Borduin CM, Schaeffer CM, Heiblum N.	2009	juvenile sexual offenders	Mean 14	48	Global Severity Index (GSI) of the Brief Symptom Inventory (BSI); Revised Behavior Problem Checklist (RBPC); Family Adaptability and Cohesion Evaluation Scales II (FACES-II); 13-item Missouri Peer Relations Inventory (MPRI); Youth reports of the Self Report Delinquency Scale (SRD);
25	Brown SA, D'Amico EJ, McCarthy DM, Tapert SF.	2001	adolescents treated for alcohol and drug problems	14-18	166	Customary drinking and drug use record (CDDR); Hollingshead classification system; Health problems composite index
26	Burleson JA, Kaminer Y, Goldston DB, Haberek R	2005	Adolescent Alcohol Abuse or Alcohol Dependence Disorder	12-18 years; average 15.9	177	Suicidal Ideation Questionnaire (SIQ-JR); Teen Addiction Severity Index (T-ASI); Diagnostic interview scale for children (self-administered Voice-DISC)
27	Burleson JA, Kaminer Y.	2005	adolescents with substance use disorder	13-18	88	Situational confidence questionnaire (SCQ); Diagnostic interview schedule for children (DISC-C)
28	Cornelius JR, Douaihy A, Bukstein OG, Daley DC, Wood SD, Kelly TM, Salloum IM.	2011	adolescents with alcohol use disorder (AUD) and major depressive disorder (MDD)	15-20	75	Hamilton rating scale for depression (HAM-D-27); Beck Depression Index (BDI); Timeline follow-back method (TLFB);

Table 3 | Continued

ID	Authors	Year	Population	Age	N	Effect measures
29	D'Amico EJ, Ellickson PL, Wagner EF, Turrisi R, Fromme K, Ghosh-Dastidar B,Longshore DL, McCaffrey DF, Montgomery MJ, Schonlau M, Wright D.	2005	adolescents with alcohol and other drug problems (AOD)	13-19	289	Perceived stress scale; Revised way of coping checklist;
30	Dembo R, Shemwell M, Guida J, Schmeidler J, Pacheco K, Seeberger W	1998	juvenile offenders	Mean 14	62	SCL-90-R;
31	Dembo R, Wothke W, Livingston S, Schmeidler J	2002	juvenile offenders	11-18; mean 14.5	278	SCL-90-R;
32	Dennis M, Godley S H, Diamond G, Tims F M, Babor T, Donaldson J, Liddle H, Titus J C, Kaminer Y, Webb C, Hamilton N, Funk R	2004	adolescent cannabis users	15-16	600	GAIN;
33	Dennis M, Titus JC, Diamond G, Donaldson J, Godley SH, Tims FM, Webb C, KaminerY, Babor T, Roebuck MC, Godley MD, Hamilton N, Liddle H, Scott CK; C. Y. T.Steering Committee.	2002	cannabis dependence or abuse in adolescents	12-18	600	Drug Abuse Treatment Cost Analysis Program (DATCAP); Global Appraisal of Individual Needs (GAIN); Adolescent Reasons for Quitting (ARFQ); Family Environment Scale (FES); Friends, Family and Self (FFS); Adolescent Relapse Coping Questionnaire (ARCQ); SCID II personality questionnaire (SPQ); Dimensions of Temperament Revised (DOTS); Child Behavioral Checklist (CBCL)
34	Gil AG, Wagner EF, Tubman JG	2004	juvenile offenders	14-19	213	Timeline follow-back interview (TLFB); Problem recognition questionnaire (PRQ)
35	Glisson C, Schoenwald SK, Hemmelgarn A, Green P, Dukes D, Armstrong KS, Chapman JE	2010	delinquent youth	9-17	child behavior checklist (CBCL)	Child behavior checklist (CBCL)

Table 3 | Continued

ID	Authors	Year	Population	Age	N	Effect measures
36	Godley SH, Garner BR, Passetti LL, Funk RR, Dennis ML, Godley MD.	2010	adolescents with substance use disorder	12-18	320	GAIN Substance Problem Scale (SPS); DATCAP;
37	Godley SH, Hedges K, Hunter B.	2011	adolescents with substance use	10-18	2141	GAIN;
38	Hall JA, Smith DC, Easton SD, An H, Williams JK, Godley SH, Jang M	2008	youth in outpatient treatment for substance abuse	12-18	404	GAIN;
39	Harold GT, Kerr DC, Van Ryzin M, Degarmo DS, Rhoades KA, Leve LD.	2013	adolescent girls in juvenile justice system	13-17	166	Brief symptom inventory (BSI);
40	Henderson CE, Dakof GA, Greenbaum PE, Liddle HA.	2010	adolescent drug abuse and delinquency	12-17	378	Personal Experience Inventory (PEI); Timeline follow-back method (TLFB); Diagnostic interview schedule for children, second edition (DISC-2); Family environment scale (FEI);
41	Henggeler SW, Halliday-Boykins CA, Cunningham PB, Randall J, Shapiro SB, Chapman JE.	2006	juvenile offenders meeting criteria for substance abuse or dependence	12-17	161	Form 90 based on TLFB; Self reported delinquency scale (SRD); Child behavior checklist (CBCL);
42	Henggeler SW, McCart MR, Cunningham PB, Chapman JE.	2012	youth substance abuse and criminal behavior	12-17	104	Form 90/TLFB; Self reported delinquency scale (SRD);
43	Henggeler SW, Melton GB, Brondino MJ, Scherer DG, Hanley JH	1997	violent and chronic juvenile offenders	11-17	155	Global severity index (GSI) of the BSI; Revised problem behavior checklist (RBPC); Self-report delinquency scale (SRD); Family adaptability and cohesion evaluation scales (FACES-III); Family assessment measure (FAM-III); Parent version monitoring index; Adolescent version monitoring index; 13-item Missouri peer relations inventory (MPRI); 14-item parent peer conformity inventory (PPCI);

Table 3 | Continued

ID	Authors	Year	Population	Age	N	Effect measures
44	Henggeler SW, Melton GB, Smith LA	1992	serious juvenile offenders	Mean 15,2	84	SRD (Self Report Delinquency Scale); FACES-III (The Family Adaptability and Cohesion Evaluation Scales); MPRI (Missouri Peer Relations Inventory); RBPC (Revised Behavior Problem Checklist); SCL-90 (Self Report Symptom Checklist); SCS-CBC;
45	Henggeler SW, Pickrel SG, Brondino MJ.	1999	substance-abusing and -dependent delinquent adolescents	12-17	118	Personal experience inventory (PEI); Self-report delinquency scale (SRD)
46	Hogue A, Dauber S, Stambaugh LF, Cecero JJ, Liddle HA.	2006	substance-abusing adolescents	average 15.5	100	TLFB; CBCL; YSR;
47	Hogue A, Henderson CE, Dauber S, Barajas PC, Fried A, Liddle HA.	2008	adolescent substance use and related behavior problems	13-17	136	Timeline follow-back (TLFB); Personal experience inventory (PEI); CBCL; YSR;
48	Hunter SB, Ramchand R, Griffin BA, Suttorp MJ, McCaffrey D, Morral A.	2012	adolescent substance use	n/a	2751	GAIN; Emotional problem scale (EPS); Illegal activities scale (IAS); Past month substance problem scale (SPS-GAIN); Substance frequency scale (SFS-GAIN);
49	Kaminer Y, Burleson JA, Goldberger R	2002	adolescent substance-abusers	13-18	88	T-ASI; DISC-C; Structural clinical interview for the DSM (SCID-II); Revised dimensions of temperament survey (DOTES-R);
50	Kaminer Y, Burleson JA.	2008	adolescents with substance use disorder	13-18	88	T-ASI; DOTES-R;

Table 3 | Continued

ID	Authors	Year	Population	Age	N	Effect measures
51	Keiley MK.	2007	incarcerated adolescents	13-18	73	Caregiver CBCL; Youth self-report (YSR); Coping inventory of stressful situations (CISS); Parental bonding instrument (PBI); Inventory of parent and peer attachment (IPPA);
52	Killeen TK, McRae-Clark AL, Waldrop AE, Upadhyaya H, Brady KT.	2012	adolescents with marijuana use disorders	12-18	31	Composite international diagnostic interview (CIDI); Teen Addiction Severity Index (T-ASI); Timeline Follow-back (TLFB); Marijuana craving questionnaire (MCQ-12); Barrett Impulsivity Scale (BIS-II-A);
53	Latimer WW, Winters KC, D'Zurilla T, Nichols M	2003	adolescent drug abusers	12-18	43	Diagnostic interview for children and adolescents (DICA-IV); Adolescent diagnostic interview revised (ADI-R)-multiaxal interview; Personal experience inventory (PEI); Family assessment measure (FAM); Social problem solving inventory (SPSI); Motivational learning questionnaire (MSLQ); Client personal history questionnaire (CPHHQ);
54	Letourneau EJ, Henggeler SW, Borduin CM, Schewe PA, McCart MR, Chapman JE, Saldana L	2009	juvenile sexual offenders	11-17	67	Adolescent sexual behavior inventory (ASBI); Self-report delinquency scale (SRD); Personal experience inventory (PEI); Child behavior checklist (CBCL);
55	Liddle HA, Dakof GA, Parker K, Diamond GS, Barrett K, Tejeda M	2001	clinically-referred marijuana- and alcohol-abusing adolescents	13-18	182	Adolescent grade point average (GPA); Acting out behaviors (AOB) scale; Global health pathology scale;
56	Liddle HA, Rowe CL, Dakof GA, Henderson CE, Greenbaum PE	2009	young adolescent substance abusers	11-15	83	GAIN; TLFB; POSIT; SRD; National youth survey peer delinquency scale;

Table 3 | Continued

ID	Authors	Year	Population	Age	N	Effect measures
57	Liddle HA, Rowe CL, Dakof GA, Ungaro RA, Henderson CE.	2004	adolescent substance abuse and behavioral problems	11-15	80	GAIN; Parent and adolescent interviews (CTRADA); Youth Self Report (YSR); Family Environment Scale (FES); National Youth Survey Peer Delinquency Scale; TLFB;
58	Lott DC, Jencius S.	2009	adolescent substance abuse	12-18	264	(only urine specimen)
59	Marsden J, Stillwell G, Barlow H, Boys A, Taylor C, Hunt N, Farrell M	2006	young ecstasy and cocaine users	16-22	342	Maudsley addiction Profile (MAP); Severity of dependence scale (SDS);
60	Martin G, Copeland J.	2008	adolescent cannabis users	14-19	40	TLFB; Items from GAIN; Severity of dependence scale (SDS); Stage of change questionnaire;
61	McCambridge J, Strang J.	2004	adolescent illegal drug use	16-20	200	Severity of Dependence Scale (SDS); Seven-point scale by Argyle; Drug Attitudes Scale (DAS); 12-item general health questionnaire (GHQ);
62	McGlynn AH, Hahn P, Hagan MP.	2012	juvenile offenders	12-18	518	HIT questionnaire
63	Moore SK, Marsch LA, Badger GJ, Solhkhah R, Hofstein Y.	2011	opoid-dependent adolescents	13-18	36	Youth Self Report (YSR);
64	Rigter H, Henderson CE, Pelc I, Tossmann P, Phan O, Hendriks V, Schaub M, RoweCL.	2012	adolescents with recent cannabis use disorder	13-18	450	Adolescent diagnostic interview-light (ADI-Light); TLFB;
65	Robbins MS, Feaster DJ, Horigian VE, Rohrbaugh M, Shoham V, Bachrach K, Miller M,Burlew KA, Hodgkins C, Carrion I, Vandermark N, Schindler E, Werstlein R,Szapocznik J.	2011	adolescent drug abuse	Mean 15.5	480	TLFB; Diagnostic interview schedule for children (DISC); Parenting Practices Questionnaire; Family environment scale (FES);

Table 3 | Continued

ID	Authors	Year	Population	Age	N	Effect measures
66	Rohde P, Jorgensen JS, Seeley JR, Mace DE	2004	incarcerated youth	12-25	76	YSR; Life attitudes schedule-short form (LAS-SF); Self-esteem scale; 10-item UCLA loneliness scale; Subjective probability questionnaire; 4 items created and modeled after the social adjustment scale;
67	Sawyer AM, Borduin CM	2011	serious and violent juvenile offenders	37,3	176	(no instruments, clinical records only)
68	Sealock MD, Gottfredson DC, Gallagher CA	1997	substance-abusing youthful offenders	n/a	460	Face valid alcohol (FVA) scale of SASSI; Face valid other drug (FVOD) scale of SASSI; Coping resources inventory (CRI); MEPS test;
69	Sexton T, Turner CW	2010	adolescents adjudicated for crime and sentenced to probation	13-17	917	WAJCA-RA structured interview
70	Timmons-Mitchell J, Bender MB, Kishna MA, Mitchell CC	2006	juvenile justice involved youth	Mean 15.1	93	Child Adolescent Functional Assessment Scale (CAFAS);
71	Waldron HB, Kern-Jones S, Turner CW, Peterson TR, Ozechowski TJ.	2007	treatment resistant drug-abusing adolescents	14-20	72	BDI; State-anger subscale of the state-trait anger expression inventory (STAXI); State-anger anxiety inventory (STAI); TLFB; CBCL; YSR; Conflict and cohesion subscale of FES;
72	Waldron HB, Slesnick N, Brody JL, Turner CW, Peterson TR.	2001	adolescent substance abuse	13-17	114	TLFB; POSIT; CBCL;
73	Walker DD, Stephens R, Roffman R, Demarce J, Lozano B, Towe S, Berg B.	2011	adolescent cannabis users	14-19	311	Global Appraisal of Individual Needs (GAIN-I); Marijuana Problem Inventory (MPI) (adapted from RAPI)

Table 3 | Continued

ID	Authors	Year	Population	Age	N	Effect measures
74	Winters KC, Fahnhorst T, Botzet A, Lee S, Lalone B	2012	Alcohol and/or cannabis use disorder	12-18	315	PCS: 11-item self-report scale from the Personal Experience Inventory (PEI); Adolescent Diagnostic Interview (ADI); TLFB; Stages of Change Readiness and Treatment Eagerness Scale (SOCRATES); Problem solving inventory; Child version of the Alabama Parenting Questionnaire (APQ); Treatment Services Review (TSR);
75	Winters KC, Leitten W.	2007	drug-abusing adolescents	14-17	79	Adolescent Diagnostic Interview (ADI) - substance use disorder module; TLFB; Personal consequences scale (PCS) from PEI; Treatment Services Review (TSR);
76	Chamberlain P, Leve LD, Degarmo DS	2007	Girls with serious and chronic delinquency	13-17	81	Elliott General Delinquency Scale;
77	Liddle, HA, Dakof, GA, Turner, RM, Henderson, CE, Greenbaum, PE	2008	Youth with drug abuse/dependence	Mean 15	224	DISC; PEI; TLFB;
78	Robbins MS, Szapocznik J, Dillon FR, Turner CW, Mitrani VB, Feaster DJ	2008	Substance-abusing or dependent adolescents	12-17, mean 15.6	190	DISC; TLFB; ADAD; Therapist Adherence Checklist
79	Smith DC, Hall JA, Williams JK, An H, Gotman N.	2006	Adolescent substance abuse	Mean 15.8	98	SFS and SPS scales of GAIN
80	Winters KC, Stinchfield RD, Opland E, Weller C, Latimer WW.	2000	Adolescent substance abuse	12-18	245	Drug consumption items of PEI.

Instrument suitability for evaluation of systemic family interventions

Table 4 displays the instruments ranked according to the number of systems covered.

Table 4 | Ranking of instruments according to the number of systems covered

Name instrument	# systems covered	systems						
		youth	family	peers	school	work	society	neighbors
POSIT	6	•	•	•	•	•	•	
CAFAS	6	•	•	•	•	•	•	
WAJCA-RA	6	•	•	•	•	•	•	
ADAD	6	•	•	•	•	•	•	
T-ASI	6	•	•	•	•	•	•	
CTRADA	5	•	•	•	•		•	
ADI	5	•	•	•	•		•	
GAIN	5	•	•		•	•	•	
PEI	4	•	•	•	•			
MAAS	4	•	•	•	•			
FES	4	•	•	•	•			
CPHHQ	3			•	•			•
FFS	3	•	•	•				
SCQ	3	•		•		•		
SRD	2	•					•	
SCL-90-R	2	•	•					
BSI	2	•	•					
Hollinghead classification system	2				•	•		
IPPA	2		•	•				
CRI	2	•	•					
MAP	2	•		•				

The majority, 81 instruments, covered just one system such as the youth or the family system. These one-dimensional instruments were often used in a multi-method (i.e. a combination of self-report, parent-report, court records, urine-analysis, etc.) assessment battery of instruments. Thirteen instruments covered two, three or four systems. We identified eight instruments, which covered five or more systems and which therefore were considered potentially suitable for comprehensive evaluation of systemic family interventions.

Detailed information on these eight instruments was searched and is highlighted below. It has to be noted that available information per instrument (e.g. number of items, example questions, domain names, most recent versions of the instrument, type of administration, etc) strongly differed.

The **Adolescent Drug Abuse Diagnosis (ADAD**) [30] is a multidimensional instrument to evaluate adolescent substance use [31] administered in a structured interview. It covers nine problem areas: medical, school, employment, social relations, family and background relations, psychological, legal, alcohol use, and drug use [32]. Example questions are "How would you rate your overall physical health?", "How many days in the past 30 have your been absent (from school)?" and "How many months did you work fulltime in the past six months?". A patient's treatment need is assessed by the interviewer per problem area based on a 10-point rating scale with scores 0-1 (no real problem), 2-3 (slight problem, treatment probably not necessary), 4-5 (moderate problem, some treatment indication), 6-7 (considerable problem, treatment necessary), and 8-9 (extreme problem, treatment absolutely necessary) [32]. The instrument consists of 150 items and is based on the Addiction Severity Index (ASI) [24]. There is also a European version of the instrument, the European Adolescent Assessment Dialogue (EuroADAD). Its aim is to "describe, communicate and compare young clients over borders of countries and institutions." [33]

The **Adolescent Diagnostic Interview (ADI)** [34] originated in the 1980's as a project "to address measurement gaps in the alcohol-drug field" [35]. It is a tool to measure substance use disorders in adolescents "…organized around *DSM–III–R* criteria for psychoactive substance use disorders." [34]. In the literature a version based on DSM-IV criteria is also mentioned [36]. The instrument is administered in a structural interview setting. Substance use of the adolescent is assessed based on two main sections with each two subsections: clinical (sociodemographics, psychosocial stressors, substance use frequency and duration, alcohol symptoms, cannabis symptoms, other substance symptoms and level functioning) and appendix (orientation and memory screen) [34]. Example items are "Which drugs have you used five or more times in your life?", "How many times do you think that you have used (this drug/each drug) in the past 6 months?", "Have you ever continuously felt like crying for several days in a row?" [36]. A computer-based version is available for self-assessment [34].

The **Child Adolescent Functional Assessment Scale (CAFAS)** "…assesses the degree of impairment in functioning in children and adolescents secondary to emotional, behavioral, or substance use problems" [37]. The instrument originally included seven scales, of which five evaluated the functioning of the youth and two scales assessed the environment of the youth [37]. The five youth scales were role performance, thinking, behavior towards self and others, moods/emotions, and substance use [38]. The two environment scales were basic needs and family/social support. The scales subsequently have been changed and expanded to 8 youth and 2 caregiver scales: school, home, community, behavior towards others, moods, self-harm, substance use, and thinking (youth) and material needs, and social support (caregiver) [38]. The different subscales include items of four severity levels (i.e. severe, moderate, mild, and minimal or no impairment) [37]. The assessor determines the level of problems of the patient per subscale. He first considers the items of the most severe level, checks whether these items apply and if not progresses towards the lesser symptom levels until an item of the current severity level applies to the patient [37]. Then scores of 30, 20, 10 and 0 are applied to severity levels severe, moderate, mild and minimal respectively such that an overall severity rating is generated. Overall ratings range from 0 to 240 with higher scores indicating higher severity [30].

The **Global Appraisal of Individual Needs (GAIN)** questionnaire [39] is a collection of related instruments that are gathered under the umbrella of GAIN using an identical format. The most recent version of the questionnaire has been adapted for use in adults as well as adolescents. The GAIN is an assessment measure, which can be used in several settings and populations such as inpatient, outpatient short- or long-term treatment evaluation, legal programs or school-based programs [40]. It assesses eight domains: background, substance use, physical health, risk behaviors, mental health, environment, legal, and vocational. Example items of the GAIN are "During the past 90 days, on how many days were you in foster care?", "When was the last time, if ever, you used...any kind of alcohol?", and "What was the most (drinks/joints/etc.) you had in one day?" [41].

The **Problem Oriented Screening Instrument for Teenagers (POSIT)** is a screening instrument for adolescents with substance use disorder, which was designed as a component of the Adolescent Assessment/Referral System (AARS) [42]. It "is designed to flag those functional areas, if any, where a problem MAY exist that requires further assessment and perhaps treatment." [42]. The instrument addresses ten functional domains: substance use/abuse, physical

health status, mental health status, family relations, peer relations, educational status, vocational status, social skills, leisure and recreation, and aggressive behavior and delinquency. The POSIT includes 139 items, which can be answered with yes or no [42]. Per domain, items can be grouped into three categories: general purpose items, general purpose age-related items, and red flag items [42]. Each affirmative response to a general purpose item counts as one point towards the total functional domain score [42]. The same holds for general purpose age-related items, but these are only relevant for specific age groups of respondents (below or above 16 years) [42]. Red flag items indicate the need for treatment once one of these items is answered positively [42]. Example items of the POSIT are "Do you get into trouble because you use drugs or alcohol at school?", "Do your parents or guardians argue a lot?", and "Have you ever been told you are hyperactive?" [42].

The **Teen Addiction Severity Index (T-ASI)** [43] is the adolescent version of the ASI [24]. The instrument assesses seven dimensions of functioning (i.e. alcohol and drug use, school status, employment-support status, family relationships, legal status, peer-social relationships, and psychiatric status) [43]. The T-ASI is intended for use in adolescents with substance use disorder aged between 12 and 19 years [43]. Example items of the T-ASI are "What chemicals have you used in the past month?", "School days spent in detention or any other measures taken for disciplinary reasons last month. (Principal's or school counselor's office.)", and "How long was your longest period of employment during the past year?" [44]. Responses are rated on a 5-point scale [43]. A revised version of the T-ASI, the T-ASI-2 has been developed in 2008. This concerns a version of the instrument, which is self-administered via computer or telephone and contains additional domains [45].

The **WAJCA-RA structured interview** is a risk assessment tool for juvenile offenders developed by the Washington State Institute for Public Policy in collaboration with the juvenile courts [46]. It was designed to identify risk and protective factors in the following domains: criminal history, school, use of free time, employment, relationships, family, alcohol and drugs, mental health, attitudes, social skills, progress on community supervision, progress while confined [46]. Example items of the WAJCA-RA are "Violence/anger: Reports of displaying a weapon, fighting, threatening people, violent outbursts, violent temper, fire starting, animal cruelty, destructiveness, volatility, intense

reactions.", "Runaways or times kicked out of home", and "Number of weeks of longest period of employment" [46].

The **Parent and adolescent interview CTRADA** that was used by [47] was not considered a common instrument but institution-specific interview as no references could be retrieved from neither literature nor the Internet. The instrument therefore could not be further considered or assessed.

Instrument suitability for use in CUA

Hence seven instruments remained for further consideration. The frequency of use of each of these instruments in the identified studies is presented in Table 5. Furthermore, Table 6 illustrates an evaluation of the instruments for suitability for use in CUA and use of preference elicitation techniques. When our feasibility characteristics were applied to the seven instruments, three instruments (POSIT, WAJCA, ADI) were excluded due to the number of domains exceeding ten, and one instrument (GAIN) was excluded due to reasons of practicability (i.e. number of items exceeding the maximum of 500 and completing time exceeding one hour). It was noted that a short version of the GAIN (Global Appraisal of Individual Needs Short Screener, GAIN-SS) is available as well [48]. However, based on its goals of screening, use for clinical staff with limited experience or periodic measurement [48], this instrument is considered too restricted for the purpose of this study. The remaining three instruments (CAFAS, T-ASI and Euro-ADAD) were considered candidates for use in CUA. One instrument (CAFAS) was considered slightly less suitable due to reasons of accessibility (i.e. concerning a paid instrument as opposed to freely available online versions of other instruments). For the remaining two instruments (T-ASI and the Euro-ADAD) only limited information on psychometric properties could be obtained. It needs noting that the T-ASI and Euro-ADAD are related as they are both based on the ASI adult instrument [33, 43]. Psychometric properties of this 'predecessor' have been judged satisfactory [24, 49-52]. To our knowledge Euro-ADAD is more frequently used in Europe, whereas T-ASI is more commonly used the United States.

Two psychometric studies with small sample sizes were identified for the T-ASI [43, 53] and one study [33] with a larger sample size was identified for the Euro-ADAD. Frequency of use was slightly favorable for the T-ASI compared to the Euro-ADAD as the instrument was used four times in the studies identified in this systematic review, whereas the Euro-ADAD was used in no more than

one study. These differences were not considered sufficient to justify favoring either of the instruments over the other. Hence, the T-ASI and Euro-ADAD were considered to have equal potential suitability for the comprehensive evaluation of systemic family interventions in a health economic framework.

Table 5 | Frequency of instrument use

Instrument name	# of papers which used this measure
Global Appraisal of Individual Needs (GAIN)	13
Teen Addiction Severity Index (T-ASI)	4
Adolescent Diagnostic Interview (ADI)	4
Problem Oriented Screening Instrument for Teenagers (POSIT)	2
Child Adolescent Functional Assessment Scale (CAFAS)	1
Washington Association of Juvenile Court 6Administrators - Risk Assessment (WAJCA-RA)	1
Adolescent Drug Abuse Diagnosis (Euro-ADAD)	1

Table 6 | Evaluation of multidimensional instruments covering five or more systems

Selection criterion	T-ASI	Euro-ADAD	CAFAS	GAIN	POSIT	WAJCA-RA	ADI
1) feasibility discrete choice experiment: - # domains<10	7	8	8 (+2 optional caregiver domains)	8	10	12	12
2) practicability of assessment in clients: - # items < 500	154	150	165	1606			
- tta < 60 min.	30-45	45-55	10	60-120			
3) accessibility: price and availability of digital version.	Digital version available free of charge	Digital version available free of charge	Available via paid online system (price $78-$400)				
4) psychometric properties [validity (content validity, construct validity, sensitivity to change); reliability (test/ retest reliability, internal consistency)] 6)	*Kaminer, 1991:* "…interrater reliability is very good; indeed, it equals or surpasses that found for most interview instruments…""Although further refinement needs to be made with respect to the Family Relationship ScaIe, the other individual scale ratings and the overall rating agreement is quite high, underscoring the reliability of the T-ASI." (N=25) *Kaminer, 1993*: "…1) T-ASI discriminated between psychoactive substance use disorders (PSUD) and non-PSUD in adolescents within a group of psychiatric inpatients; 2) T-ASI substance use, psychiatric status, family function, and school status scores were related to external criteria; 3) there was specificity in these relations with external criteria…support for the valid psychometric properties of the T-ASI….All results should be interpreted in the context of the small sample size (N=25)."	*Czobor, 2011:* Good test-retest reliability (Pearson's r >= 0.8 for most individual domains) and internal consistency ("Based on Cronbach's coefficient alpha the internal consistency/reliability of the instrument was 0.50… when we applied an extension of the internal consistency/ reliability measure for multidimensional case, we found that the internal consistency/reliability was satisfactory."); good criterion, convergent and discriminant construct validity. (N=632)					
5) frequency of use in the field: - number of times used in studies	4	1					

Discussion and conclusions

The objective of this systematic literature review was to identify existing instruments in the field of adolescent delinquency and substance use, which cover the relevant domains of systemic family interventions. The instruments were appraised based on characteristics relevant for use in economic evaluations such as brevity, accessibility, psychometric properties etc. Euro-ADAD and T-ASI showed favorable characteristics in relation to the criteria for a comprehensive outcome measure, covering multiple relevant systems and being suitable for obtaining preference weights. Both instruments lack preference weights for the outcomes, at present. Attaining these (as a potential next step) would facilitate calculating 'utility scores' as common in economic evaluations. Furthermore, the results of the current study may inform future efforts towards standardized and comprehensive core outcome sets as defined by the COMET initiative [54]. The study may be seen as a preparatory step towards a full COMET effort to standardizing the QALY approach to include broader effects.

Some limitations of this study must be noted. First, given our focus on published research up to 2013, we may have missed out on very recent developments in this field. In the Netherlands, for instance, a new, comprehensive instrument for measuring substance abuse in adolescents is being developed, called the MATE-Y [55], which includes nine modules each containing several domains. Yet, up to today there have not yet been publications on the MATE in the field of youth/adolescents. But similar developments may be ongoing elsewhere. Second, we have not investigated the possibility of constructing a new measure by combining different measures into one composite measure. Though this may be a limitation of this paper, we considered it a necessary first step to identify the instruments currently available in the field for direct use. This may also help to highlight the relevant domains to include in a newly developed instrument. With our approach, we were able to identify two instruments as most promising candidates to use in comprehensive evaluations of systemic family interventions. Neither instrument is currently considered 'gold standard' in practice. Furthermore, as common for systematic reviews, the results from the current study are based on a limited selection of databases within a limited timeframe. Yet the number of screened and identified articles was extensive and we assume that the consultation of an even larger number of databases would not have yielded significant differences in results. Also, the characteristics for further selection of the instruments were necessarily arbitrary and guided by our goal of selecting one or more instruments suitable to be used to attain societal preference

weights and be used in economic evaluations in the long term. We realize that the suitability criterion of a maximum of 500 questions/1 hour of completion time may be rather high when considering the busy clinical practice and ongoing evaluation of patient progress. Furthermore, had we considered different or more broad characteristics, additional instruments might have been found suitable. For example, one could think of shortening existing longer instruments first and then proceeding towards steps of attaining societal preference weights. In the light of limited time, this was not considered feasible in the current study.

Notwithstanding these limitations, our review revealed two promising, currently used instruments, which may be made suitable for inclusion in economic evaluations of systemic family interventions: the Euro-ADAD and T-ASI. To make these instruments suitable for health economic evaluations, first of all, more detailed investigation is necessary of their validity, feasibility and comprehensiveness. Current information on this is scarce, yet needed. Moreover, using these instruments in health economic evaluations will require important next steps. In particular, preference weights would need to be derived for the different states described by the instrument, like those available for health-related utility measures such as the EQ-5D. This is possible through preference elicitation techniques, such as discrete choice experiments or time-trade-off techniques, ultimately leading to 'utility scores', which can be attached to the different 'states' described by the instrument.

Intriguing questions in this context relate to who should indicate the state a person is in and who should provide the values for the different possible states (i.e., whose preferences count). In line with many guidelines for health-economic evaluations [11, 56], and in line with the broad aim of systemic family interventions, one could ask 'patients' to provide self-reports based on one of the identified multidimensional instruments. The value attached to this state could then be based on preferences obtained in the general population. This would provide 'societal weights' for the broad outcomes of systemic family interventions. These societal weights could thus be attached to the state a person indicates him- or herself to be in on the multidimensional instrument, thus leading to an overall utility score. Given the broad range of outcomes, including effects incurred by others than the patient or even his family (e.g., a safe neighborhood), the score thus relates to a preference ordering over states that include the effects on more than the patient alone. This may be an additional reason for opting for general public preferences. However, whether the general

public is the appropriate source (rather than e.g. decision makers or health care professionals) must be further assessed and discussed, as well as their ability to appropriately weight such diverse outcomes. The more fundamental question is whether these scores would count as 'utilities' or rather as multi-criteria decision weights.

Other relevant issues in developing a multidimensional utility measure of systemic family interventions may be the diversity and hierarchy of treatment effects. As mentioned earlier, a comprehensive measure would include health as well as non-health effects and would also include both the effects on the patient himself and society as a whole. Obviously, these different effects may be interrelated. Moreover, some observable effects may be considered to be intermediate effects, whereas others may be final outcomes. Related to this point, there may be short-term and long-term effects, which can be important. Hence, in the construction of such a preference-based measure, good care needs to be taken of the possible interaction of the effects.

One may argue that an alternative route to finding an appropriate outcome measure could be to use existing measures in the field of economic evaluation, most notably QALY measures. To our knowledge, so far there have been only a few studies on the validity of preference-weighted health-related quality of life instruments in an adult population of substance abusers [57, 58]. There have been two studies on the degree to which common preference-weighted measures of quality of life (e.g. QWB-SA, SF-12) correlate with substance use severity [58, 59]. Whereas the first study provides evidence for insufficient coverage of all disease dimensions in substance use disorder [58], the second study does suggest moderate to good correlation between quality of life measures and substance use severity measures [59]. In order to verify these results and determine whether the proposed instruments add value in the field of delinquency and substance abuse in adolescents, further research on the suitability and potential of the quality adjusted life year (QALY) measure in this population is recommended.

Keeping these alternatives in mind, further research on the instruments highlighted in the current paper, specifically on the attachment of societal preference weights could bring evaluation of mental health interventions for delinquent and substance abusing adolescents closer to the standard methodology in health economic evaluations of curative medical interventions. Both identified instruments appear suitable and broad enough to capture the

effects of family interventions in substance abusing and delinquent adolescents in such CUA. Adding societal preference weights to one of these instruments will create an instrument, which combines the advantage of the specificity of a disorder-specific instrument with compliance with common methodology of health economic evaluations and captures the broad effects relevant to mental health interventions. CUAs of these interventions can then be performed based on a broad and specific measure that includes several systems/dimensions and at the same time acknowledges the relative value that society attaches to improvements in these diverse systems. Though performing CUAs in the field of substance abuse and delinquency in adolescents remains a challenging task, this paper attempted to contribute to confronting one of the major issues in that context: finding a suitable outcome measure.

References

[1] Cottrell D, Boston P: Practitioner review: The effectiveness of systemic family therapy for children and adolescents. *J Child Psychol Psychiatry* 2002, 43(5): 573-586.

[2] Randall J, Cunningham P B: Multisystemic therapy: a treatment for violent substance-abusing and substance-dependent juvenile offenders. *Addict Behav* 2003, 28(9): 1731-1739.

[3] Brysbaert M: *Psychologie.* Gent: Academia Press; 2009.

[4] Hendriks V, van der Schee E, Blanken P: Treatment of adolescents with a cannabis use disorder: Main findings of a randomized controlled trial comparing multidimensional family therapy and cognitive behavioral therapy in The Netherlands. *Drug Alcohol Depend* 2011.

[5] Henggeler S W, Sheidow A J: Empirically supported family-based treatments for conduct disorder and delinquency in adolescents. *J Marital Fam Ther* 2002, 38(1): 30-58.

[6] French M T, Zavala S K, McCollister K E, Waldron H B, Turner C W, Ozechowski T J: Cost-effectiveness analysis of four interventions for adolescents with a substance use disorder. *J Subst Abuse Treat* 2008, 34(3): 272-281.

[7] Borduin C M, Mann B J, Cone L T, Henggeler S W, Fucci B R, Blaske D M, et al: Multisystemic treatment of serious juvenile offenders: long-term prevention of criminality and violence. *J Consult Clin Psychol* 1995, 63(4): 569-578.

[8] Henggeler S W, Clingempeel W G, Brondino M J, Pickrel S G: Four-year follow-up of multisystemic therapy with substance-abusing and substance-dependent juvenile offenders. *J Am Acad Child Adolesc Psychiatry* 2002, 41(7): 868-874.

[9] Liddle H A, Rowe C L, Dakof G A, Henderson C E, Greenbaum P E: Multidimensional family therapy for young adolescent substance abuse: twelve-month outcomes of a randomized controlled trial. *J Consult Clin Psychol* 2009, 77(1): 12-25.

[10] Sexton T, Turner C W: The effectiveness of functional family therapy for youth with behavioral problems in a community practice setting. Journal of family psychology: JFP : journal of the Division of Family Psychology of the American Psychological Association (Division 43) 2010, 24(3): 339-348.

[11] College voor Zorgverzekeringen: *Dutch guidelines for pharmacoeconomic research, revised edition*. Diemen: College voor Zorgverzekeringen; 2006.

[12] National Institute for Health and Care Excellence: *Guide to the methods of technology appraisal*; 2013.

[13] Goorden M, Van Der Schee E, Hendriks VM, Hakkaart-van Roijen L: Cost-effectiveness of multidimensional family therapy compared to cognitive behavioral therapy for adolescents with a cannabis use disorder: Data from a randomized controlled trial. *Drug Alcohol Depend* 2016, May 1;162:154-61.

[14] Brazier J: Is the EQ–5D fit for purpose in mental health?. *Br J Psychiatry* 2010, *197*(5): 348-349.

[15] Knapp M, Mangalore R: The trouble with QALYs... . *Epidemiol Psychiatr Soc* 2007, 16(4): 289-293.

[16] Brooks R, Group E: EuroQol: the current state of play. *Health policy* 1996, *37*(1): 53-72.

[17] Al-Janabi H, Flynn T N, Coast J: Estimation of a preference-based carer experience scale. *Med Decis Making 2011*, *31*(3): 458-468.

[18] Al-Janabi H, Coast J, Flynn T N: What do people value when they provide unpaid care for an older person? A meta-ethnography with interview follow-up. *Soc Sci Med 2008*, *67*(1): 111-121.

[19] Sanders GD, Neumann PJ, Basu A, Brock DW, Feeny D, Krahn M, Kuntz KM, Meltzer DO, Owens DK, Prosser LA, Salomon JA. Recommendations for conduct, methodological practices, and reporting of cost-effectiveness analyses: second panel on cost-effectiveness in health and medicine. Jama. 2016 Sep 13;316(10):1093-103.

[20] Goorden M, Schawo S J, Bouwmans-Frijters C A, van der Schee E, Hendriks V M, Hakkaart-van Roijen L. The cost-effectiveness of family/family-based therapy for treatment of externalizing disorders, substance use disorders and delinquency: a systematic review. *BMC psychiatry 2016*, *16*(1): 237.

[21] Jofre-Bonet M, Sindelar J L: Creating an aggregate outcome index: cost-effectiveness analysis of substance abuse treatment. *J Behav Health Serv Res* 2004, 31(3): 229-241.

[22] Deas D, Thomas S E: An overview of controlled studies of adolescent substance abuse treatment. *Am J Addict* 2001, 10(2): 178-189.

[23] Hogue A, Liddle H A: Family-based treatment for adolescent substance abuse: controlled trials and new horizons in services research. *J Fam Ther* 2009, 31(2), 126-154.

[24] McLellan A T, Luborsky L, Woody G E, O'Brien C P: An improved diagnostic evaluation instrument for substance abuse patients. The Addiction Severity Index. *J Nerv Ment Dis* 1980, 168(1): 26-33.

[25] Bronfenbrenner U: Contexts of child rearing: Problems and prospects. *Am Psychol* 1979, *34*(10): 844.

[26] Hinton W J, Sheperis C, Sims P: Family-based approaches to juvenile delinquency: A review of the literature. *Fam J Alex Va* 2003, *11*(2): 167-173.

[27] Moher D, Liberati A, Tetzlaff J, Altman D G, Prisma Group: Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS med* 2009, *6*(7): e1000097.

[28] Ryan M, Gerard K: Using discrete choice experiments to value health care programmes: current practice and future research reflections. *Appl Health Econ Health Policy* 2003, 2(1): 55--64.

[29] Witt J, Scott A, Osborne R. Designing choice experiments with many attributes. An application to setting priorities for orthopaedic waiting lists. *J Health Econ* 2009, 18: 681--696.

[30] Friedman A S, Utada A: A method for diagnosing and planning the treatment of adolescent drug abusers (the Adolescent Drug Abuse Diagnosis [ADAD] instrument). *J Drug Educ* 1989, 19(4): 285-312.

[31] Chinet L, Plancherel B, Bolognini M, Holzer L, Halfon O: Adolescent substance-use assessment: methodological issues in the use of the Adolescent Drug Abuse Diagnosis (ADAD). *Subst Use Misuse* 2007, 42(10): 1505-1525.

[32] European Monitoring Centre for Drugs and Drug Addiction. Adolescent Drug Abuse Diagnosis [http://www.emcdda.europa.eu/html.cfm/index3530EN.html]

[33] Czobor P, Bacskai E, Oberg D, Haack M J, Gerevich J: The European Adolescent Assessment Dialogue (EuroADAD): a psychometric evaluation. *Eur Addict Res* 2011, 17(6): 302-315.

[34] Winters K C, Stinchfield R D, Fulkerson J, Henly G A: Measuring Alcohol and Cannabis Use Disorders in an Adolescent Clinical Sample. *Psychol Addict Behav* 1993, 7(3): 185-196.

[35] Winters K C, Stinchfield R D, Henly G A.: Further validation of new scales measuring adolescent alcohol and other drug abuse. *J Stud Alcohol* 1993, 54(5): 534-541.

[36] European Monitoring Centre for Drugs and Drug Addiction. Adolescent Diagnostic Interview (ADI) [http://www.emcdda.europa.eu/attachements.cfm/att_4137_EN_tadi.pdf.]

[37] Hodges K, Wong M M: Psychometric Characteristics of a Multidimensional Measure to Assess Impairment: The Child and Adolescent Functional Assessment Scale. *J CHILD FAM STUD* 1996, 5(4), 445-467.

[38] Multi-Health Systems Inc. CAFAS - Child and Adolescent Functional Assessment Scale. [http://www.mhs.com/product.aspx?gr=cli&prod=cafas&id=overview]

[39] Dennis M, Godley S H, Diamond G, Tims F M, Babor T, Donaldson J, et al: The Cannabis Youth Treatment (CYT) Study: main findings from two randomized trials. *J SUBST ABUSE TREAT* 2004, 27(3): 197-213.

[40] Chestnut Health Systems. Gain Global Appraisal of Individual Needs. [http://www.gaincc.org/about-gain/]

[41] GAIN Coordinating Center. . GAIN - Global Appraisal of Individual Needs. [http://gaincc.org/_data/files/Instruments%20and%20Reports/Instruments%20Manuals/GAIN-I_GAIN-M90_Walkthrough_5.6.pdf.]

[42] Rahdert E R: The Adolescent Assessment/Referral System Manual 1991.

[43] Kaminer Y, Bukstein O, Tarter R E. The Teen-Addiction Severity Index: rationale and reliability. *Int J Addict* 1991, 26(2), 219-226.

[44] Kaminer, Y: Teen Addiction Severity Index (T-ASI). [http://www.emcdda.europa.eu/attachements.cfm/att_4008_EN_tt-asi.pdf]

[45] Brodey B B, McMullin D, Kaminer Y, Winters K C, Mosshart E, Rosen C S, et al: Psychometric characteristics of the Teen Addiction Severity Index-Two (T-ASI-2). Substance abuse: official publication of the Association for Medical Education and Research in Substance Abuse 2008, 29(2): 19-32.

[46] Washington State Institute for Public Policy. Washington State Juvenile Court Assessment Manual Version 1999 2.0. [http://www.wsipp.wa.gov/rptfiles/99-01-0000.pdf]

[47] Liddle H A, Rowe C L, Dakof G A, Ungaro R A, Henderson C E: Early intervention for adolescent substance abuse: pretreatment to posttreatment outcomes of a randomized clinical trial comparing multidimensional family therapy and peer group treatment. *J Psychoactive Drugs* 2004, 36(1): 49-63.

[48] Chestnut Health Systems. . GAIN-SS. [http://gaincc.org/gainss]

[49] Hendriks V M, Kaplan C D, van Limbeek J, Geerlings P: The Addiction Severity Index: reliability and validity in a Dutch addict population. *J Subst Abuse Treat* 1989, 6(2): 133-141.

[50] Leonhard C, Mulvey K, Gastfriend D R, Shwartz M: The Addiction Severity Index: a field study of internal consistency and validity. *J Subst Abuse Treat* 2000, 18(2): 129-135.

[51] McLellan A T, Luborsky L, Cacciola J, Griffith J, Evans F, Barr H L, et al: New data from the Addiction Severity Index. Reliability and validity in three centers. *J Nerv Ment Dis* 1985, 173(7): 412-423.

[52] Zanis D A, McLellan A T, Cnaan R A, Randall M: Reliability and validity of the Addiction Severity Index with a homeless sample. *J Subst Abuse Treat* 1994, 11(6): 541-548.

[53] Kaminer Y, Wagner E, Plummer B, Seifer R: Validation of the Teen Addiction Severity Index (T-ASI). American Academy of Psychiatrists in Alcoholism and Addictions 1993, 2(3): 250-254.

[54] COMET (Core Outcome Measures in Effectiveness Trials) Initiative. http://www.comet-initiative.org/.

[55] Schippers G, Broekman T: MATE-Y 2.1a - handleiding en protocol voor de MATE-jeugd. Nijmegen: Beta Boeken; 2013.

[56] Gold M, Siegel J, Russel L, Weinstein M: Cost-Effectiveness in Health and Medicine. New York: Oxford University Press, USA; 1996.

[57] van der Zanden B P, Dijkgraaf M G, Blanken P, de Borgie C A, van Ree J M, van den Brink W: Validity of the EQ-5D as a generic health outcome instrument in a heroin-dependent population. *Drug Alcohol Depend* 2006, 82(2): 111-118.

[58] Pyne J M, French M, McCollister K, Tripathi S, Rapp R, Booth B: Preference-weighted health-related quality of life measures and substance use disorder severity. Addict 2008, 103(8): 1320-9; discussion 1330-2.

[59] Pyne J M, Tripathi S, French M, McCollister K, Rapp R C, Booth B M: Longitudinal association of preference-weighted health-related quality of life measures and substance use disorder outcomes. Addict 2011, 106(3): 507-515.

References (systematic review results)

1. Armelius, B. A., & Andreassen, T. H. (2007). Cognitive-behavioral treatment for antisocial behavior in youth in residential treatment. The Cochrane database of systematic reviews, (4)(4), CD005650.

2. Arnold, E. M., Kirk, R. S., Roberts, A. C., Griffith, D. P., Meadows, K., & Julian, J. (2003). Treatment of incarcerated, sexually-abused adolescent females: an outcome study. Journal of Child Sexual Abuse, 12(1), 123-139.

3. Azzopardi, P. S., Kennedy, E. C., Patton, G. C., Power, R., Roseby, R. D., Sawyer, S. M., et al. (2013). The quality of health research for young Indigenous Australians: systematic review. The Medical journal of Australia, 199(1), 57-63.

4. Baldwin, S. A., Christian, S., Berkeljon, A., & Shadish, W. R. (2012). The effects of family therapies for adolescent delinquency and substance abuse: a meta-analysis. Journal of marital and family therapy, 38(1), 281-304.

5. Borduin, C. M. (1999). Multisystemic treatment of criminality and violence in adolescents. Journal of the American Academy of Child and Adolescent Psychiatry, 38(3), 242-249.

6. Borduin, C. M. (1999). Multisystemic treatment of criminality and violence in adolescents. Journal of the American Academy of Child and Adolescent Psychiatry, 38(3), 242-249.

7. Borduin, C. M., Schaeffer, C. M., & Heiblum, N. (2009). A randomized clinical trial of multisystemic therapy with juvenile sexual offenders: effects on youth social ecology and criminal activity. Journal of consulting and clinical psychology, 77(1), 26-37.

8. Brown, S. A., & D'Amico, E. J. (2003). Outcomes of alcohol treatment for adolescents. Recent developments in alcoholism : an official publication of the American Medical Society on Alcoholism, the Research Society on Alcoholism, and the National Council on Alcoholism, 16, 289-312.

9. Brown, S. A., D'Amico, E. J., McCarthy, D. M., & Tapert, S. F. (2001). Four-year outcomes from adolescent alcohol and drug treatment. Journal of studies on alcohol, 62(3), 381-388.

10. Burleson, J. A., & Kaminer, Y. (2005). Self-efficacy as a predictor of treatment outcome in adolescent substance use disorders. Addictive Behaviors, 30(9), 1751-1764.

11. Chamberlain, P., Leve, L. D., & Degarmo, D. S. (2007). Multidimensional treatment foster care for girls in the juvenile justice system: 2-year follow-up of a randomized clinical trial. Journal of consulting and clinical psychology, 75(1), 187-193.

12. Cornelius, J. R., Douaihy, A., Bukstein, O. G., Daley, D. C., Wood, S. D., Kelly, T. M., et al. (2011). Evaluation of cognitive behavioral therapy/motivational enhancement therapy (CBT/MET) in a treatment trial of comorbid MDD/AUD adolescents. Addictive Behaviors, 36(8), 843-848.

13. Cottrell, D., & Boston, P. (2002). Practitioner review: The effectiveness of systemic family therapy for children and adolescents. Journal of child psychology and psychiatry, and allied disciplines, 43(5), 573-586.

14. Curtis, N. M., Ronan, K. R., & Borduin, C. M. (2004). Multisystemic treatment: a meta-analysis of outcome studies. Journal of family psychology : JFP : journal of the Division of Family Psychology of the American Psychological Association (Division 43), 18(3), 411-419.

15. D'Amico, E. J., Ellickson, P. L., Wagner, E. F., Turrisi, R., Fromme, K., Ghosh-Dastidar, B., et al. (2005). Developmental considerations for substance use interventions from middle school through college. Alcoholism, Clinical and Experimental Research, 29(3), 474-483.

16. Deas, D. (2008). Evidence-based treatments for alcohol use disorders in adolescents. Pediatrics, 121 Suppl 4, S348-54.

17. Deas, D., & Thomas, S. E. (2001). An overview of controlled studies of adolescent substance abuse treatment. The American Journal on Addictions / American Academy of Psychiatrists in Alcoholism and Addictions, 10(2), 178-189.

18. Dembo, R., Shemwell, M., Guida, J., Schmeidler, J., Pacheco, K., & Seeberger, W. (1998).

A Longitudinal Study of the Impact of a Family Empowerment Intervention on Juvenile Offender Psychosocial Functioning: A First Assessment

. Journal of Child & Adolescent Substance Abuse, 8(1), 15-54.

19. Dembo, R., Wothke, W., Livingston, S., & Schmeidler, J. (2002). The impact of a family empowerment intervention on juvenile offender heavy drinking: a latent growth model analysis. Substance use & misuse, 37(11), 1359-1390.

20. Dennis, M., Godley, S. H., Diamond, G., Tims, F. M., Babor, T., Donaldson, J., et al. (2004). The Cannabis Youth Treatment (CYT) Study: main findings from two randomized trials. Journal of substance abuse treatment, 27(3), 197-213.

21. Dennis, M., Titus, J. C., Diamond, G., Donaldson, J., Godley, S. H., Tims, F. M., et al. (2002). The Cannabis Youth Treatment (CYT) experiment: rationale, study design and analysis plans. Addiction (Abingdon, England), 97 Suppl 1, 16-34.

22. Diamond, G., & Josephson, A. (2005). Family-based treatment research: a 10-year update. Journal of the American Academy of Child and Adolescent Psychiatry, 44(9), 872-887.

23. Ferguson, L. M., & Wormith, J. S. (2012). A Meta-Analysis of Moral Reconation Therapy. International Journal of Offender Therapy and Comparative Criminology.

24. Gil, A. G., Wagner, E. F., & Tubman, J. G. (2004). Culturally sensitive substance abuse intervention for Hispanic and African American adolescents: empirical examples from the Alcohol Treatment Targeting Adolescents in Need (ATTAIN) Project. Addiction (Abingdon, England), 99 Suppl 2, 140-150.

25. Glisson, C., Schoenwald, S. K., Hemmelgarn, A., Green, P., Dukes, D., Armstrong, K. S., et al. (2010). Randomized trial of MST and ARC in a two-level evidence-based treatment implementation strategy. Journal of consulting and clinical psychology, 78(4), 537-550.

26. Godley, S. H., Garner, B. R., Passetti, L. L., Funk, R. R., Dennis, M. L., & Godley, M. D. (2010). Adolescent outpatient treatment and continuing care: main findings from a randomized clinical trial. Drug and alcohol dependence, 110(1-2), 44-54.

27. Godley, S. H., Hedges, K., & Hunter, B. (2011). Gender and racial differences in treatment process and outcome among participants in the adolescent community reinforcement approach. Psychology of addictive behaviors : journal of the Society of Psychologists in Addictive Behaviors, 25(1), 143-154.

28. Hall, J. A., Smith, D. C., Easton, S. D., An, H., Williams, J. K., Godley, S. H., et al. (2008). Substance abuse treatment with rural adolescents: issues and outcomes. Journal of psychoactive drugs, 40(1), 109-120.

29. Harold, G. T., Kerr, D. C., Van Ryzin, M., Degarmo, D. S., Rhoades, K. A., & Leve, L. D. (2013). Depressive Symptom Trajectories Among Girls in the Juvenile Justice System: 24-month Outcomes of an RCT of Multidimensional Treatment Foster Care. Prevention science : the official journal of the Society for Prevention Research.

30. Henderson, C. E., Dakof, G. A., Greenbaum, P. E., & Liddle, H. A. (2010). Effectiveness of multidimensional family therapy with higher severity substance-abusing adolescents: report from two randomized controlled trials. Journal of consulting and clinical psychology, 78(6), 885-897.

31. Henggeler, S. W., Halliday-Boykins, C. A., Cunningham, P. B., Randall, J., Shapiro, S. B., & Chapman, J. E. (2006). Juvenile drug court: enhancing outcomes by integrating evidence-based treatments. Journal of consulting and clinical psychology, 74(1), 42-54.

32. Henggeler, S. W., McCart, M. R., Cunningham, P. B., & Chapman, J. E. (2012). Enhancing the effectiveness of juvenile drug courts by integrating evidence-based practices. Journal of consulting and clinical psychology, 80(2), 264-275.

33. Henggeler, S. W., Melton, G. B., Brondino, M. J., Scherer, D. G., & Hanley, J. H. (1997). Multisystemic therapy with violent and chronic juvenile offenders and their families: the role of treatment fidelity in successful dissemination. Journal of consulting and clinical psychology, 65(5), 821-833.

34. Henggeler, S. W., Melton, G. B., & Smith, L. A. (1992). Family preservation using multisystemic therapy: an effective alternative to incarcerating serious juvenile offenders. Journal of consulting and clinical psychology, 60(6), 953-961.

35. Henggeler, S. W., Pickrel, S. G., & Brondino, M. J. (1999). Multisystemic treatment of substance-abusing and dependent delinquents: outcomes, treatment fidelity, and transportability. Mental health services research, 1(3), 171-184.

36. Henggeler, S. W., & Sheidow, A. J. (2012). Empirically supported family-based treatments for conduct disorder and delinquency in adolescents. Journal of marital and family therapy, 38(1), 30-58.

37. Henggeler, S. W., & Sheidow, A. J. (2003). Conduct disorder and delinquency. Journal of marital and family therapy, 29(4), 505-522.

38. Hogue, A., Dauber, S., Stambaugh, L. F., Cecero, J. J., & Liddle, H. A. (2006). Early therapeutic alliance and treatment outcome in individual and family therapy for adolescent behavior problems. Journal of consulting and clinical psychology, 74(1), 121-129.

39. Hogue, A., Henderson, C. E., Dauber, S., Barajas, P. C., Fried, A., & Liddle, H. A. (2008). Treatment adherence, competence, and outcome in individual and family therapy for adolescent behavior problems. Journal of consulting and clinical psychology, 76(4), 544-555.

40. Hogue, A., & Liddle, H. A. (2009). Family-based treatment for adolescent substance abuse: controlled trials and new horizons in services research. Journal of Family Therapy, 31(2), 126-154.

41. Hunter, S. B., Ramchand, R., Griffin, B. A., Suttorp, M. J., McCaffrey, D., & Morral, A. (2012). The effectiveness of community-based delivery of an evidence-based treatment for adolescent substance use. Journal of substance abuse treatment, 43(2), 211-220.

42. Kaminer, Y., & Burleson, J. A. (2008). Does temperament moderate treatment response in adolescent substance use disorders?. Substance abuse : official publication of the Association for Medical Education and Research in Substance Abuse, 29(2), 89-95.

43. Kaminer, Y., Burleson, J. A., & Goldberger, R. (2002). Cognitive-behavioral coping skills and psychoeducation therapies for adolescent substance abuse. The Journal of nervous and mental disease, 190(11), 737-745.

44. Kaminer, Y., Burleson, J. A., Goldston, D. B., & Burke, R. H. (2006). Suicidal ideation among adolescents with alcohol use disorders during treatment and aftercare. The American Journal on Addictions / American Academy of Psychiatrists in Alcoholism and Addictions, 15 Suppl 1, 43-49.

45. Keiley, M. K. (2007). Multiple-family group intervention for incarcerated adolescents and their families: a pilot project. Journal of marital and family therapy, 33(1), 106-124.

46. Killeen, T. K., McRae-Clark, A. L., Waldrop, A. E., Upadhyaya, H., & Brady, K. T. (2012). Contingency management in community programs treating adolescent substance abuse: a feasibility study. Journal of child and adolescent psychiatric nursing : official publication of the Association of Child and Adolescent Psychiatric Nurses, Inc, 25(1), 33-41.

47. Latimer, W. W., Winters, K. C., D'Zurilla, T., & Nichols, M. (2003). Integrated family and cognitive-behavioral therapy for adolescent substance abusers: a stage I efficacy study. Drug and alcohol dependence, 71(3), 303-317.

48. Letourneau, E. J., Henggeler, S. W., Borduin, C. M., Schewe, P. A., McCart, M. R., Chapman, J. E., et al. (2009). Multisystemic therapy for juvenile sexual offenders: 1-year results from a randomized effectiveness trial. Journal of family psychology : JFP : journal of the Division of Family Psychology of the American Psychological Association (Division 43), 23(1), 89-102.

49. Liddle, H. A., Dakof, G. A., Parker, K., Diamond, G. S., Barrett, K., & Tejeda, M. (2001). Multidimensional family therapy for adolescent drug abuse: results of a randomized clinical trial. The American Journal of Drug and Alcohol Abuse, 27(4), 651-688.

50. Liddle, H. A., Dakof, G. A., Turner, R. M., Henderson, C. E., & Greenbaum, P. E. (2008). Treating adolescent drug abuse: a randomized trial comparing multidimensional family therapy and cognitive behavior therapy. Addiction (Abingdon, England), 103(10), 1660-1670.

51. Liddle, H. A., Rowe, C. L., Dakof, G. A., Henderson, C. E., & Greenbaum, P. E. (2009). Multidimensional family therapy for young adolescent substance abuse: twelve-month outcomes of a randomized controlled trial. Journal of consulting and clinical psychology, 77(1), 12-25.

52. Liddle, H. A., Rowe, C. L., Dakof, G. A., Ungaro, R. A., & Henderson, C. E. (2004). Early intervention for adolescent substance abuse: pretreatment to posttreatment outcomes of a randomized clinical trial comparing multidimensional family therapy and peer group treatment. Journal of psychoactive drugs, 36(1), 49-63.

53. Littell, J. H., Popa, M., & Forsythe, B. (2005). Multisystemic Therapy for social, emotional, and behavioral problems in youth aged 10-17. The Cochrane database of systematic reviews, (4)(4), CD004797.

54. Lott, D. C., & Jencius, S. (2009). Effectiveness of very low-cost contingency management in a community adolescent treatment program. Drug and alcohol dependence, 102(1-3), 162-165.

55. Marsden, J., Stillwell, G., Barlow, H., Boys, A., Taylor, C., Hunt, N., et al. (2006). An evaluation of a brief motivational intervention among young ecstasy and cocaine users: no effect on substance and alcohol use outcomes. Addiction (Abingdon, England), 101(7), 1014-1026.

56. Martin, G., & Copeland, J. (2008). The adolescent cannabis check-up: randomized trial of a brief intervention for young cannabis users. Journal of substance abuse treatment, 34(4), 407-414.

57. McCambridge, J., & Strang, J. (2004). The efficacy of single-session motivational interviewing in reducing drug consumption and perceptions of drug-related risk and harm among young people: results from a multi-site cluster randomized trial. Addiction (Abingdon, England), 99(1), 39-52.

58. McGlynn, A. H., Hahn, P., & Hagan, M. P. (2012). The Effect of a Cognitive Treatment Program for Male and Female Juvenile Offenders. International Journal of Offender Therapy and Comparative Criminology.

59. Moore, S. K., Marsch, L. A., Badger, G. J., Solhkhah, R., & Hofstein, Y. (2011). Improvement in psychopathology among opioid-dependent adolescents during behavioral-pharmacological treatment. Journal of addiction medicine, 5(4), 264-271.

60. Randall, J., & Cunningham, P. B. (2003). Multisystemic therapy: a treatment for violent substance-abusing and substance-dependent juvenile offenders. Addictive Behaviors, 28(9), 1731-1739.

61. Rigter, H., Henderson, C. E., Pelc, I., Tossmann, P., Phan, O., Hendriks, V., et al. (2013). Multidimensional family therapy lowers the rate of cannabis dependence in adolescents: a randomised controlled trial in Western European outpatient settings. Drug and alcohol dependence, 130(1-3), 85-93.

62. Robbins, M. S., Feaster, D. J., Horigian, V. E., Rohrbaugh, M., Shoham, V., Bachrach, K., et al. (2011). Brief strategic family therapy versus treatment as usual: results of a multisite randomized trial for substance using adolescents. Journal of consulting and clinical psychology, 79(6), 713-727.

63. Robbins, M. S., Szapocznik, J., Dillon, F. R., Turner, C. W., Mitrani, V. B., & Feaster, D. J. (2008). The efficacy of structural ecosystems therapy with drug-abusing/dependent African American and Hispanic American adolescents. Journal of family psychology : JFP : journal of the Division of Family Psychology of the American Psychological Association (Division 43), 22(1), 51-61.

64. Rohde, P., Jorgensen, J. S., Seeley, J. R., & Mace, D. E. (2004). Pilot evaluation of the Coping Course: a cognitive-behavioral intervention to enhance coping skills in incarcerated youth. Journal of the American Academy of Child and Adolescent Psychiatry, 43(6), 669-676.

65. Sealock, M., Gottfredson, D., & Gallagher, C. (1997).

Drug Treatment for Juvenile Offenders: Some Good and Bad News . Journal of Research in Crime and Delinquency, 34(2), 210-236.

66. Sexton, T., & Turner, C. W. (2010). The effectiveness of functional family therapy for youth with behavioral problems in a community practice setting. Journal of family psychology : JFP : journal of the Division of Family Psychology of the American Psychological Association (Division 43), 24(3), 339-348.

67. Smith, D. C., Hall, J. A., Williams, J. K., An, H., & Gotman, N. (2006). Comparative efficacy of family and group treatment for adolescent substance abuse. The American Journal on Addictions / American Academy of Psychiatrists in Alcoholism and Addictions, 15 Suppl 1, 131-136.

68. Tanner-Smith, E. E., Wilson, S. J., & Lipsey, M. W. (2013). The comparative effectiveness of outpatient treatment for adolescent substance abuse: a meta-analysis. Journal of substance abuse treatment, 44(2), 145-158.

69. Timmons-Mitchell, J., Bender, M. B., Kishna, M. A., & Mitchell, C. C. (2006). An independent effectiveness trial of multisystemic therapy with juvenile justice youth. Journal of clinical child and adolescent psychology : the official journal for the Society of Clinical Child and Adolescent Psychology, American Psychological Association, Division 53, 35(2), 227-236.

70. Tripodi, S. J., Bender, K., Litschge, C., & Vaughn, M. G. (2010). Interventions for reducing adolescent alcohol abuse: a meta-analytic review. Archives of Pediatrics & Adolescent Medicine, 164(1), 85-91.

71. Waldron, H. B., & Kaminer, Y. (2004). On the learning curve: the emerging evidence supporting cognitive-behavioral therapies for adolescent substance abuse. Addiction (Abingdon, England), 99 Suppl 2, 93-105.

72. Waldron, H. B., Kern-Jones, S., Turner, C. W., Peterson, T. R., & Ozechowski, T. J. (2007). Engaging resistant adolescents in drug abuse treatment. Journal of substance abuse treatment, 32(2), 133-142.

73. Waldron, H. B., Slesnick, N., Brody, J. L., Turner, C. W., & Peterson, T. R. (2001). Treatment outcomes for adolescent substance abuse at 4- and 7-month assessments. Journal of consulting and clinical psychology, 69(5), 802-813.

74. Waldron, H. B., & Turner, C. W. (2008). Evidence-based psychosocial treatments for adolescent substance abuse. Journal of clinical child and adolescent psychology : the official journal for the Society of Clinical Child and Adolescent Psychology, American Psychological Association, Division 53, 37(1), 238-261.

75. Walker, D. D., Stephens, R., Roffman, R., Demarce, J., Lozano, B., Towe, S., et al. (2011). Randomized controlled trial of motivational enhancement therapy with nontreatment-seeking adolescent cannabis users: a further test of the teen marijuana check-up. Psychology of addictive behaviors : journal of the Society of Psychologists in Addictive Behaviors, 25(3), 474-484.

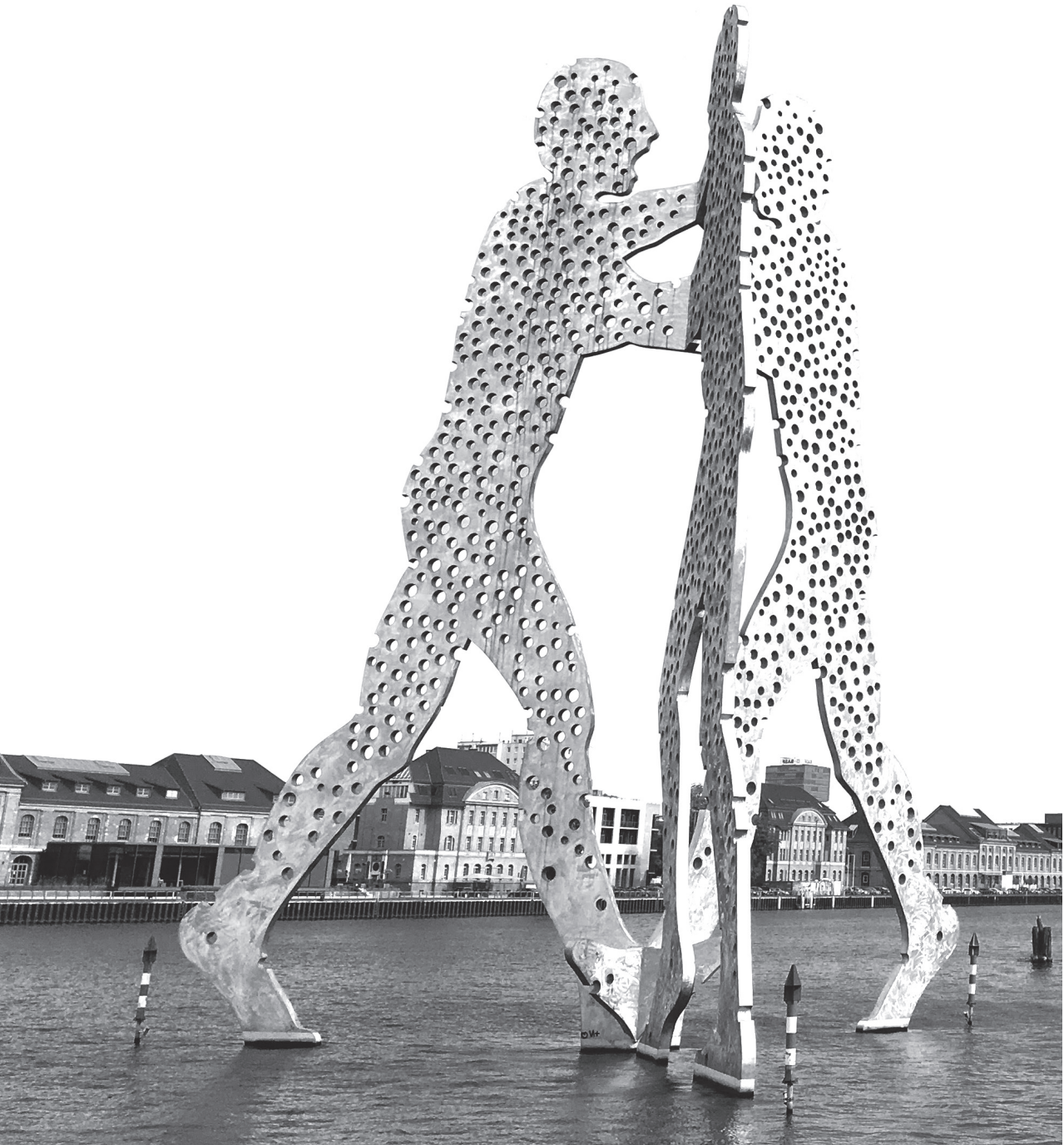
76. Walker, D. F., McGovern, S. K., Poey, E. L., & Otis, K. E. (2004). Treatment effectiveness for male adolescent sexual offenders: a meta-analysis and review. Journal of Child Sexual Abuse, 13(3-4), 281-293.

77. Winters, K. C., Fahnhorst, T., Botzet, A., Lee, S., & Lalone, B. (2012). Brief intervention for drug-abusing adolescents in a school setting: outcomes and mediating factors. Journal of substance abuse treatment, 42(3), 279-288.

78. Winters, K. C., & Leitten, W. (2007). Brief intervention for drug-abusing adolescents in a school setting. Psychology of addictive behaviors : journal of the Society of Psychologists in Addictive Behaviors, 21(2), 249-254.

79. Winters, K. C., Stinchfield, R. D., Opland, E., Weller, C., & Latimer, W. W. (2000). The effectiveness of the Minnesota Model approach in the treatment of adolescent drug abusers. Addiction (Abingdon, England), 95(4), 601-612.

80. Woolfenden, S. R., Williams, K., & Peat, J. (2001). Family and parenting interventions in children and adolescents with conduct disorder and delinquency aged 10-17. The Cochrane database of systematic reviews, (2)(2), CD003015.



Chapter 8

Obtaining Preference Scores for an Abbreviated Self-completion Version of the Teen-Addiction Severity Index (ASC T-ASI) to Value Therapy Outcomes of Systemic Family Interventions: a Discrete Choice Experiment

Based on Schawo, S., Hoefman, R., Reckers-Droog, V.T., van de Wetering, L., Kaminer, Y., Brouwer, W.B.F., Hakkaart-van Roijen, L.

Submitted to Health and Quality of Life Outcomes.

Introduction

Economic evaluations in health care often take the form of cost-utility analysis, in which outcomes are captured in terms of quality-adjusted life-years (QALYs) [1,2], as measured with generic instruments like the EQ-5D [3] or SF-6D [4]. This implicitly reflects that many curative health care interventions primarily aim to improve health and longevity of patients. However, in certain health care sectors the aim of interventions may not (primarily) be to improve health, but to improve broader aspects of quality of life than health alone. These outcomes may be captured insufficiently by existing generic health-related quality of life instruments used to calculate QALYs [5, 6]. This issue is gaining attention, for example, in the area of elderly care, where broader measures like the ICECAP-O and ASCOT have been developed [7, 8]. These measures capture broader life domains than health and are suitable for use in economic evaluations. In other areas, including mental healthcare and addiction-related treatments, such generic preference-based outcome measures are still lacking [8-12]. This issue of appropriate and comprehensive outcome measures, preference-based and suitable for use in economic evaluations, is highly relevant in the context of mental health interventions aimed at youths and systemic family interventions in particular. These interventions are intended to have broad effects (e.g. related to substance use, family interactions, interaction with peers, and performance at school) exceeding the domain of health. If not appropriately identified, measured and valued, such broader effects may fall outside the scope of economic evaluations, risking wrong conclusions about value for money of interventions and, potentially, undesirable (that is, welfare lowering) decisions concerning their reimbursement.

The relevance of this issue is emphasized by the fact that systemic family interventions for adolescents with problems of substance use and/or delinquency are increasingly subject of economic evaluations [13]. However, existing studies are limited in quality and comparability as settings, design and outcome measures vary extensively [13]. The application of economic evaluation in the field of systemic family interventions is hampered by the lack of preference-based instruments that are validated, sensitive and feasible to use and capture all relevant benefits. Systemic family interventions are explicitly directed at improving interactions between the adolescent patient and surrounding systems, and are often used in the context of substance abuse and delinquency [14-16]. Aims of such interventions are diverse and include improvements in family relations, peer interactions, achievements at work or school, and reduction of

substance use and criminal activity [15; 17-19]. In a meta-analysis evaluating the effectiveness of outpatient substance abuse treatments for adolescents, systemic family interventions were found to be effective in the treatment of substance abuse [20]. Given that these interventions typically are intensive and costly [17; 21; 22], economic evaluations are important, also to inform reimbursement decisions. This requires validated, broad multidimensional preference-based instruments that capture the relevant effects of such interventions.

In order to fill this gap, a recent systematic review of the effectiveness literature on systemic family interventions identified existing instruments, which measure benefits beyond health-related quality of life [12]. While no preference-based instruments were found, the Teen-Addiction Severity Index (T-ASI) [23] was identified as a multidimensional instrument that captures the main relevant life domains affected by these interventions. Although preference scores for this instrument were lacking, it was considered suitable for adaptation into a preference-based measure for use in economic evaluations of systemic family interventions [12].

The original T-ASI is a relatively long semi-structured interview that measures symptoms of adolescent substance use based on seven domains and five levels of problem severity. The instrument is not a self-report instrument but completed by a therapist together with the patient. Some questions are directed at the patient while others ask the therapist to provide his or her judgment. In order to make the instrument suitable for use in economic evaluations, in which patients commonly report their own situation using a self-complete descriptive system, we created an abbreviated, patient-completed version of the T-ASI, the ASC T-ASI [24; see appendix A and B]. This abbreviated instrument was based on the main patient-reported questions from all domains of the T-ASI, reflecting the functioning of the patient as judged by him or herself. The ASC T-ASI is a broad outcome measure, suitable for self-completion. After designing the instrument, two studies were performed. One study validated the ASC T-ASI, with favorable results [24]. The second derived societal preference scores for the ASC T-ASI using a discrete choice experiment (DCE). This second study is presented in the current paper.

Methods

Questionnaire

Data was obtained from an existing online panel. People who signed up for the panel were invited to participate in this study. Those who accepted the invitation were informed about the purpose of the study and how anonymity of participants was guaranteed. They were informed that participation in the study was voluntary and could be stopped at any time, in which case that the data they had provided up to that point would be discarded. By submitting their response at the end of the questionnaire they provided consent for the use of their data for the stated purposes of the study. Participants received no financial compensation.

The Dutch translation of the ASC T-ASI [24; Appendix B] forms the basis for the current study and the preference-based measure. In this study, we used a DCE to obtain societal preference weights for all domains and levels of the ASC T-ASI instrument. A professional Dutch translation agency advised us in formulating the instructions of the discrete choice tasks based on reading level B1.

Pilot and main data were collected with a questionnaire that was distributed online by a professional survey company. As the common source of health state valuations is the general public [1, 2], we elicited preferences for different outcomes described with the ASC T-ASI in a sample representative of the general adult population in the Netherlands in terms of age (18-65 years), gender, and level of education. Before respondents completed the questionnaire, they were informed about the background of the study, the target population of adolescents with problems of substance abuse and/or delinquency, and the attributes and levels of the DCE. Furthermore, an outline of the questionnaire, instructions on the type of questions and a privacy statement were provided. The questionnaire of pilot and main study comprised four parts. Part one included questions about demographics of the respondent. Part two consisted of the DCE tasks. Part three stated questions about the feasibility and readability of the DCE tasks. Part four consisted of questions on current health status of the respondent.

Discrete choice experiment

DCEs are frequently used to inform policy decisions in health care [25-27]. In such experiments individuals are confronted with a series of choice tasks. The DCE methodology is based on McFadden's random utility theory [28] and assumes that an individual, when confronted with a choice task that consists of

n alternatives with a fixed number of attributes and attribute levels, will choose the option that maximizes his or her utility.

Choice task

The current study used choice tasks with two unlabeled alternatives (A and B) reflecting a state of an adolescent described by the seven attributes (substance use, school, work, family, social relationships, justice, and mental health) and five attribute levels (ranging from 'no problem' to 'very large problem') of the ASC T-ASI. Unlike the case in common outcome measures, where respondents are asked to value states for themselves, given the specific nature of the ASC T-ASI, respondents here were asked to choose between the alternatives based on what they believed would be best for the adolescent (hence the question was formulated in the third person). In this way, we obtained societal preferences for the different situations of the adolescent described with the ASC T-ASI instrument. Appendix C presents an example of one of the choice tasks, as presented to the respondents.

Design pilot study

Between 11 and 16 December 2013, we collected pilot data from a sample (n=106) representative of the general adult population in the Netherlands in terms of age (18-65 years), gender, and education level. The pilot study had two main objectives. First, we collected information on the attributes and levels that could be used for the development of an efficient design for the main study. Second, we obtained information concerning the feasibility and readability of the DCE tasks.

A dummy-coded multinomial logit (MNL) model with fixed priors was chosen to build a D-efficient design in NGENE version 1.1.2. As higher problem levels were logically expected to be associated with lower preference scores in all seven attributes, priors were fixed at 0.04, 0.03, 0.02, and 0.01 for 'no problem' to 'large problem', respectively. 'Very large problem' was set as the base case attribute level for all attributes. The design included 50 rows and respondents were randomly assigned to one of five blocks, resulting in ten choice sets per respondent. In each of the presented choice tasks, respondents were asked to imagine an adolescent with problems of substance use and/or delinquency and to choose the alternative, which they considered to reflect the best scenario for the adolescent. To force respondents to choose between one of the provided alternatives, no opt-out was provided. To examine left-right bias in respondent choices, an alternative-specific constant was added to the model.

Two control tasks were included to identify respondents who responded inconsistently. The first task was a dominated choice scenario, with one alternative indicating less problems in all domains. The second control task presented respondents with a mirrored version of a choice task they had already answered earlier on in the DCE. Respondents who answered at least one of the two control questions inconsistently were excluded. Expected completion time was 12 minutes based on the mean completion time determined by test runs by two independent researchers. Responders who completed the questionnaire in less than a third of this time (<4 minutes) were considered to be 'speeders' and were excluded from the analysis.

Design main study

Between 7 and 13 March 2014, we collected data from a sample (n=1,500) representative of the general adult population in the Netherlands in terms of age (18-65 years), gender, and education level. Based on the results of the pilot study, the design of the main study was slightly adapted. The number of questions per respondent was reduced from ten to eight and color-coding was applied to the choice tasks to visually emphasize the differences in problem severity between attribute levels. A D-efficient design with 40 rows and five blocks was created by applying normally distributed Bayesian priors from the pilot study and using 1000 Halton draws. The attribute levels were dummy-coded. Respondents were randomly assigned to one of five blocks with eight choice tasks each plus two control tasks. In line with the pilot study, respondents were excluded from the analyses when they were considered 'speeders' or answered at least one of the two control questions incorrectly.

Model specification

The main data were analyzed by first applying an MNL model and stepwise extending this model towards a panel mixed multinomial logit (MMNL) model. Whereas for the MNL model the assumption holds that all variables need to be independent and identically distributed (IID assumption), this assumption does not apply to the panel MMNL model [29]. The panel MMNL model hence allows for interdependency of observations (which e.g. may occur when respondents answer several choice tasks) and heterogeneity in respondent preferences. Consequently, within the panel MMNL model, utility variation, which would otherwise enter into the error component of the MNL model is explicitly modeled and reflected in the parameter estimates [29, 30]. Model fit was evaluated based on log likelihood ratio (LR) tests.

When extending the MNL model towards a panel MMNL model, several steps were taken. First, an unrestricted dummy-coded MNL model with an alternative-specific constant was estimated. No evidence for left-right bias was found, and hence the constant was excluded from the model. Next, we investigated various model specifications with random parameters to allow for heterogeneity in respondent preferences. Making all parameters random was not feasible technically due to limitations on computer capacity, so stepwise parameters that indicated the strongest heterogeneity, i.e. with the highest standard error, were added as random parameters and model fit was evaluated based on LR tests. As a final step, we verified whether collapsing attribute levels two and three ('fairly large problem' and 'large problem') or three and four ('large problem' and 'very large problem') would improve model fit. However, these modifications did not lead to an improvement based on LR. Hence, a panel MMNL with fourteen random and fourteen fixed parameters was chosen as the final model. Standard deviations were derived based on Cholesky decomposition. The analyses were performed in NLOGIT (version 5).

ASC T-ASI preference scores

To estimate the coefficients for the fourteen random parameters, bootstrapping using 10,000 hypothetical individuals from a normal distribution using the population level estimates of the MMNL was applied and individual-specific parameters were derived. Individual-specific parameters for each of the attributes and levels were averaged. The averages of the random parameters and estimates of the beta coefficients of the non-random parameters from the MMNL model were rescaled to a 0-1 scale to provide an ASC T-ASI tariff set. A score of 0 refers to the worst state with very large problems in all of the ASC T-ASI domains, while a score of 1 refers to no problems in any of the domains.

Results

Pilot study

The pilot study included 106 respondents (after the exclusion of 'speeders' and respondents who answered at least one of the control questions inconsistently). The number of excluded respondents was unknown as the survey company directly excluded these. Analysis of the dummy-coded MNL model of the pilot study showed that the intercept and the five coefficients of the attributes 'substance use' and 'justice' were significant at the 5% level; the remaining coefficients were not significant at the 5% level. In addition, coefficient values of 'substance use' and 'justice' were generally larger than those of the remaining

attributes. Based on significance and size of the coefficients we concluded that the weight of the attributes 'substance use' and 'justice' could be considered higher than the weight of the remaining attributes in the main study. As an indication of these possible differences in preferences, we set the priors for the design of the main study at 0.8; 0.7; 0.6; 0.5 for the attributes 'substance use' and 'justice' and at 0.4; 0.3; 0.2; 0,1 for the remaining attributes.

Table 1 | Results of multinomial logit regression model (pilot study)

Attribute	Level	β coefficient	Standard error
Alternative specific constant		.17213**	.07364
Substance use	No problem	-.25458	.15787
	Slight problem	-.33906**	.14340
	Fairly large problem	-.29451*	.16178
	Large problem	-.43473**	.16930
	Very large problem	Base	
School	No problem	.02548	.14697
	Slight problem	-.01061	.14727
	Fairly large problem	-.00028	.13802
	Large problem	-.11372	.13429
	Very large problem	Base	
Work	No problem	-.10070	.13417
	Slight problem	-.25961*	.13853
	Fairly large problem	.11855	.15717
	Large problem	-.20238	.15715
	Very large problem	Base	
Family	No problem	-.17069	.17254
	Slight problem	-.13104	.14296
	Fairly large problem	-.19364	.15630
	Large problem	-.24786	.15315
	Very large problem	Base	
Social relationships	No problem	-.02810	.15607
	Slight problem	.00062	.15381
	Fairly large problem	-.26423	.18211
	Large problem	-.04516	.15244
	Very large problem	Base	
Justice	No problem	-.37518**	.15755
	Slight problem	-.25657	.15922
	Fairly large problem	-.34744***	.13473
	Large problem	-.34762**	.15835
	Very large problem	Base	
Mental health	No problem	.19324	.15250
	Slight problem	.03586	.14204
	Fairly large problem	.03471	.15899
	Large problem	.09234	.14292
	Very large problem	Base	

Note: ***, **, * ==> Significance at 1%, 5%, 10% level respectively.

Concerning the feasibility and readability of the questionnaire, 56.6% (*n=60*) of respondents of the pilot study found the DCE tasks 'difficult' or 'very difficult'. The majority of respondents had difficulties choosing between the different alternatives (*n=53*). Some respondents had difficulties imagining the situation of the adolescent (*n=10*), or reading the descriptions provided with the choice tasks (*n=4*). To reduce complexity and enhance clarity for respondents in the main study, the questionnaire was slightly adapted as compared to the pilot study. The first adaptation referred to color-coding of the attribute levels from light to dark violet as to facilitate a choice by making the different levels visually more distinct. In addition, the number of choice sets was reduced from 10 to 8 per respondent, to increase the feasibility of the study and quality of the data.

Main study

The main study comprised 1,500 respondents after exclusion of 'speeders' and exclusion of 853 respondents who answered at least one of the control questions incorrectly. General respondent characteristics are displayed in Table 2.

Table 2 | Sample characteristics of main study (n=1,500)

	Sample main study Mean (SD) or % (n)	General Dutch population Mean (SD) or %
Gender (Male)	50.3 (*754*)	50.2
Age		
Female		
18-34	16.3 (*244*)	16.2
35-49	16.8 (*252*)	16.7
50-65	16.7 (*250*)	16.9
Male		
18-34	16.5 (*248*)	16.5
35-49	16.9 (*253*)	16.8
50-65	16.9 (*253*)	17.0
Educational level[a]		
Minimum Low	10.9 (164)	
Minimum Middle	58.6 (879)	
Minimum High	30.5 (457)	
Completion time (min)		
Minimum	4.00	
Maximum	23.70	
Mean	8.16	
Work		
Yes (>=36 hrs/wk)	29.3 (*439*)	
Yes (<36 hrs/wk)	27.6 (*414*)	
No	43.1 (*647*)	
Children		
Yes (<12 yrs)	13.1 (*196*)	
Yes (12-21 yrs)	18.5 (*278*)	
Yes (>21 yrs)	26.3 (*395*)	
No	50.8 (*762*)	
Subjective health (EQ-5D-3L VAS)	74.35 (14.78)	77.72 (15.19)[b]
Health-related QoL (EQ-5D-3L)	.87 (0.20)	.87 (0.18)[b]

Note. [a] Low = lower vocational and primary school, Middle = middle vocational and secondary school, High = higher vocational and academic education; [b] [32]

Respondents' distribution of age and gender was in line with the general population in the Netherlands. Mean age was 42 years (general Dutch population: 42 years) and the proportion of male respondents was 50.3% (general Dutch population: 50.2%) [31]. Completion time ranged from 4 minutes to nearly 24 minutes, with a mean completion time of 8 minutes and 13 seconds. This shorter

completing time as compared to the pilot can be explained by the reduction in the number of choice sets and the addition of color-coding. A large proportion of the respondents stated not to have paid work (43.1%), 29.3% worked 36 or more hours a week and 27.6% worked part time with an average of 19.16 hours per week (SD 9.063). Slightly more than half of respondents (50.8%) had children, of which 18.5% were between 12 and 21 years (an age group similar to the population that the discrete choice task referred to). Subjective health and health-related quality of life based on the EQ-5D and its Dutch tariffs were comparable to the values of the general Dutch population [32].

Preference scores for the ASC T-ASI domains and problem levels

The results of the panel MMNL model are presented in table 3. An overview of the coefficients for each of the attributes and attribute levels is provided.

Table 3 | Results of panel mixed multinomial logit regression model (main study)

Attribute	Level	β coefficient	Standard deviation
Substance use	No problem	2.48564***	1.63320***
	Slight problem	1.92631***	1.39210***
	Fairly large problem	.87412***	-
	Large problem	.33400***	-
	Very large problem	Base	
School	No problem	.91027***	-
	Slight problem	.86616***	.37781***
	Fairly large problem	.33925**	1.10870***
	Large problem	.25556***	-
	Very large problem	Base	
Work	No problem	1.25942***	1.08894***
	Slight problem	.81447***	-
	Fairly large problem	.58935***	-
	Large problem	.21731*	.79275***
	Very large problem	Base	
Family	No problem	1.67697***	-
	Slight problem	1.17190***	-
	Fairly large problem	.58164***	.85331***
	Large problem	.03268	.94046***
	Very large problem	Base	
Social relationships	No problem	1.26387***	.73073***
	Slight problem	1.01831***	-
	Fairly large problem	.78538***	-
	Large problem	.37928***	1.02409***
	Very large problem	Base	
Justice	No problem	2.02321***	1.53078***
	Slight problem	1.53178***	1.09865***
	Fairly large problem	.68487***	-
	Large problem	.25052**	.66371***
	Very large problem	Base	
Mental health	No problem	2.31869***	1.06997***
	Slight problem	1.95064***	-
	Fairly large problem	1.21795***	-
	Large problem	.52154***	-
	Very large problem	Base	

Note: ***, **, * ==> Significance at 1%, 5%, 10% level respectively; - ==>fixed parameters

Table 3 shows that all coefficients were positive. Hence, generally, fewer problems than the base case level (very large problems) were preferred by the respondents. The coefficients of problems with substance use, family, justice and mental health were relatively large compared to the other coefficients indicating that changes in these domains had a relatively high impact on the choice between alternative situations. Problems in the domains school, work and social relationships had a relatively low impact on the choice. All but two coefficients were significant at the 5% level. One coefficient for 'large' problems was only significant at the 10% level (with an effect of 0.019 on the tariff), and one was not significant at the 10% level (with a marginal effect of 0.001 on the tariff). This suggests that in these two cases there was no evidence for the level of a 'large problem' to be significantly different from a 'very large problem' (base level). Yet, as described above, collapsing the problem levels 'large' and 'very large' did not yield significant improvements of the model (as shown by the LR).

Table 3 also shows that all standard deviations of the random parameters were relative large and significant at the 1% level hence providing evidence for preference heterogeneity amongst respondents.

Tariff set

Table 4 presents the results of the conversion of the coefficients into preference scores per domain and problem level with the total score ranging from 0 to 1. A score of 0 refers to the worst state as defined by the instrument and a score of 1 refers to the best possible state defined by the instrument.

The use of the preference scores can be illustrated as follows. Based on table 4, an adolescent with a 'slight problem' in the domain substance use, a 'fairly large problem' in the domains school and work and 'no problem' in the domains family, social relationships, justice and mental health would be coded 2331111, which translates into a score of 0.161+0.028+0.050+0.141+0.106+0.168+0.194 =0.848.

Consistent with the coefficients presented in Table 3 and the above mentioned example it can be seen that the domains substance use, mental health, justice, and family were more influential and received more weight than the domains social relationships, work and school.

Table 4 | ASC T-ASI tariff set

Domain	Problem level	Preference Scores
Substance use	No problem	.210
	Slight problem	.161
	Fairly large problem	.073
	Large problem	.028
	Very large problem	.000
School	No problem	.076
	Slight problem	.073
	Fairly large problem	.028
	Large problem	.022
	Very large problem	.000
Work	No problem	.105
	Slight problem	.068
	Fairly large problem	.050
	Large problem	.019
	Very large problem	.000
Family	No problem	.141
	Slight problem	.098
	Fairly large problem	.049
	Large problem	.001
	Very large problem	.000
Social relationships	No problem	.106
	Slight problem	.086
	Fairly large problem	.066
	Large problem	.032
	Very large problem	.000
Justice	No problem	.168
	Slight problem	.128
	Fairly large problem	.058
	Large problem	.022
	Very large problem	.000
Mental health	No problem	.194
	Slight problem	.164
	Fairly large problem	.102
	Large problem	.044
	Very large problem	.000

Discussion

In this study, we obtained societal preference scores for the ASC T-ASI, creating a short preference-based measure suitable for use in economic evaluations of systemic family interventions. The scope of this measure is more in line with the goals of systemic interventions than currently available health-related quality of life measures, and hence enables a more meaningful interpretation of the effects of such interventions. The instrument is an adaptation of the frequently used T-ASI [23], which may contribute to its acceptance, validity and feasibility of implementation. The ASC T-ASI is a preference-based outcome measure with a self-contained scoring system with its own range and interpretation. In that respect it is similar to other recently developed broader outcome measures, like the ICECAP instruments [5, 6]. While the preference-based ASC T-ASI has clear similarities with common health-related quality of life measures, we emphasize that they are distinct. For health-related quality of life measures like the EQ-5D, a preference-score of 0 corresponds to the state 'dead' and hence represents a 'natural zero'. This is not the case for the ASC T-ASI, where a score of 0 simply refers to the most severe problems on all domains of the instrument and where 1 refers to the best situation. Combining ASC T-ASI scores with duration therefore requires a careful consideration and interpretation. Moreover, QALY tariffs represent average valuations of health states obtained by asking respondents to imagine being in these health states themselves. Here, we asked adults to value states from the perspective of the adolescent, not themselves, which is a fundamental difference. The resulting scores therefore cannot be straightforwardly transformed into or compared with QALYs. The use of the preference-based ASC T-ASI measure hence implies the loss of some of the comparability specific to CUA (as comparison is only possible between interventions which can be evaluated with the same quality of life concept or even instrument). However, it may be more informative and yield more meaningful results when performing economic evaluations of systemic family interventions where effects broader than health may be expected and allows comparisons of benefits of such interventions.

Sindelar and Jofre-Bonet [33] have earlier expressed the need for a preference-weighted instrument to perform CEAs in the context of substance abuse treatment. They presented an index score for the ASI, the adult version of the T-ASI, which differs in some domains from the T-ASI. The authors obtained index scores by asking patients and participants from a convenience sample from the general public how important treatment was or how important each

domain was. The authors introduced their method as an "intermediate step until further, more sophisticated surveys become available" [33]. The current study may be considered such a further step, providing societal preferences for the actual states the ASC T-ASI describes, based on a discrete choice experiment.

Before discussing some implications and future research, we note a number of limitations and strengths of this study. A first limitation relates to the issue of completeness, i.e. whether all relevant effects of systemic interventions are covered in the ASC T-ASI. One might argue that more dimensions could have been included in the instrument. Moreover, some of the included domains may not be relevant for all respondents. For example, the domain 'work' may only be relevant for relatively old adolescents who work or would want to work. Future research may be directed at investigating this issue further, for instance by considering conditional questions or changes in the labeling of the domains or levels.

A second limitation is related to the significance of parameters. Ideally, all parameters would be significant. Two parameters presented in table 3 (family–large problems and work – large problems) were not significant at the 5% level, although the latter was significant at the 10% level (p-value 0.089). This may indicate some inefficiencies in the design, yet the impact of the non-significant coefficients on the tariff was limited to negligible (with values of 0.019 and 0.001, respectively). Non-significance of these parameters suggests that the levels of 'large problems' and 'very large problems' may not need to be evaluated separately but could be merged for these domains. Merging however resulted in a reduced model fit. Hence, we chose to keep the levels apart, basing the tariff on all available information and differentiating between problem levels in the same way in all domains. Sample-size may have had an influence on significance levels as well.

Third, interactions between attributes and attribute levels were not explicitly modeled due to the large number of model parameters and limitations in computer capacity. This may constitute a shortcoming of the current study and may be explored in future research.

Fourth, a more general limitation, related to the design of the study, is that respondents stated to have experienced the choice tasks as complex. This was already observed in the pilot study. In the main study, we therefore decreased

the number of choice tasks from ten to eight per respondent and applied color-coding to simplify the decision process and reduce overall demands to respondents. Furthermore, to increase the probability that the choice tasks were well understood by the included respondents, speeders and respondents who did not answer the control questions correctly were excluded. Nonetheless, more than half of the included respondents still considered questions in the main study to be difficult or very difficult. When provided with possible reasons for the difficulties, 44.1% (n=661) of respondents indicated problems with making a choice between the different situations. This may be related to the inherently difficult nature of the choices in this context. Other problems were less frequently mentioned, including having trouble imagining the situations of the adolescent, trouble reading the descriptions of the choice tasks, and possible interdependencies of alcohol and drug problems with other problems. Part of the difficulties of making a choice may have been related to the fact that respondents were confronted with a forced choice without an opt out option. While this was done intentionally to avoid disturbing the utility balance of the design, lowering its efficiency, and to avoid respondents opting out due to reasons other than preferences related to the choice task [34], the absence of such an option may have increased the difficulty of the task. Including an opt-out option might have influenced our results.

Fifth, potentially related to the previous point, we excluded a substantial amount of respondents who incorrectly answered one or two control questions, in order to achieve the highest possible quality of the data for the tariff set. One could argue that excluding only those respondents who answered both control questions incorrectly would have been sufficiently cautious. Respondents who were excluded due to answering one control question incorrectly (N=717) were significantly older (44.91 vs. 42.00 years; p=0.000) and lower educated (p=0.000) than included respondents. No gender difference was observed.

Sixth, very important also for the interpretation of the presented tariffs, as indicated above, the observed scores reflect what people in the general public think is 'best' for the adolescents involved, not an indication of a preference to be in a certain state oneself. This may have added to the difficulty of the task. It also represents a crucial difference with many other preference elicitations in health care in which people normally choose for themselves. This difference between the here presented tariffs from common 'utility scores' needs emphasis. We chose this valuation approach for several reasons. Firstly, we wanted to obtain

broad societal preferences from the general public, in line with Dutch guidelines [1]. Given that the ASC T-ASI relates to adolescents, this renders obtaining preferences of people for themselves meaningless. Hence, we asked them to opt the best option for the adolescent, which yielded preferences that may be somewhat 'paternalistic'. Secondly, we assumed that preferences of adolescents actually being in these states arguably would be less useful for societal decision making, especially when 'distorted' by underlying problems like addiction and myopia (also due to the age of respondents). Moreover, such preferences would be influenced by coping and adaptation [35].

Future research could explore the important normative issue of 'whose values count' [35] in situations like these, but could also compare preferences of affected adolescents, adolescents without the specific problems described with the instrument, and those of the general public. Using preferences from non-affected adolescents may yield preferences that are more representative of those of the treated group. Moreover, arguably, such respondents might be more capable of imagining (what it means) being in the different states described with the ASC T-ASI than adults in the general population. However, whether their preferences would be (more) appropriate to use in societal decision making remains unclear. Domains like 'social relationships' could for instance receive more weight in samples consisting of adolescents, at the expense of domains like family or school.

Future research may also consider the framing of the choice task. We chose the framing of asking which situation was 'best for the adolescent', reflecting potential treatment goals of the health system, which can be different from what the adolescent would prefer. The approach taken therefore can be viewed as being aligned with societal decision-making and collective financing of interventions, at the expense of not using preferences of the treated adolescents.

Seventh, duration of states was not included as an attribute. Hence, when applying the ASC T-ASI preference-based measure in practice in combination with duration linearity of scores over time needs to be assumed.

Finally, the current study was limited to the Dutch setting. Moreover, among our respondents there was a high percentage of individuals without paid work, which may have affected our results.

A number of strengths of this study also deserve mentioning. First, we built on an existing instrument to come to the ASC T-ASI, and the first validation study of the adapted version of the instrument showed promising results [24]. This approach may lead to higher acceptance and feasibility when implementing the measure in cost-effectiveness studies or in clinical practice. The Dutch version of the ASC T-ASI is set up in reading level B1, which may have enhanced readability, facilitating self-completion also by adolescents. Second, we used a two stage-design, starting with an elaborate pilot, followed by a main study. Advantages of this approach were that adjustments to the design could be made in between the pilot and main study, enhancing the quality of the data obtained. A third strength relates to the model used in our analyses, which allowed for interdependency of observations and heterogeneity in preferences. The chosen model fits the panel data of the DCE tasks and accounts for individual differences in choice behavior, which is valuable as we assume that choices may differ between different members of society. Future research could further analyze the heterogeneity in the data. Our current aim was to obtain overall preference scores rather than to differentiate between the scores of specific groups of respondents.

The presented ASC T-ASI can be used in several ways. It may be used as an add-on instrument in future cost-effectiveness studies and clinical trials with low burden to patients due to its brevity. Also, it can be used as a stand-alone self-completion instrument to weight different changes in the situation of adolescents. Both options would provide valuable information for use in economic evaluations, in combination with the here provided tariffs. When used in combination with other cost or benefit measures in economic evaluation, overlap and double counting need to be avoided. Such overlap could occur with common measures like the EQ-5D or with cost components of economic evaluations. This, as well as the validity of the ASC T-ASI in different settings, needs to be investigated further in future research. Furthermore, though the ASC T-ASI is developed in the context of systemic family interventions, future studies may consider its application in a broader context of youth mental health interventions.

Concluding, we performed a discrete choice experiment to attain preference scores for the ASC T-ASI. Our goal was to facilitate the use of the ASC T-ASI in the context of economic evaluations, by obtaining specific preference scores for this instrument capturing the most relevant disease-specific aspects of systemic

family interventions in adolescents with problems with substance use and/or delinquency. To our knowledge, the ASC T-ASI is the first generic preference-based instrument in adolescent mental health care for which societal preference scores have been obtained that capture benefits beyond those included in the QALY. Many questions for further research were identified which exceed the scope of the current study. Nonetheless, the presented tariff hopes to provide a first step in including relevant disease-specific aspects in economic evaluations of systemic family interventions.

References

[1] The Health Care Insurance Board (CVZ). Guidelines for Pharmacoeconomic Research [in Dutch]. CVZ. Diemen; 2006.

[2] Drummond MF, Sculpher MJ, Torrance GW, O'Brien BJ, Stoddart GL. Methods for the economic evaluation of health care programmes. Oxford University Press; 2005.

[3] Brooks R, Rabin R, De Charro F. The measurement and valuation of health status using EQ-5D: a European perspective: evidence from the EuroQol BIOMED Research Programme. Springer; 2003.

[4] Brazier J, Roberts J, Deverill M. The estimation of a preference-based measure of health from the SF-36. Journal of health economics. 2002, Mar;21(2):271-92.

[5] Al-Janabi H, Flynn T, Coast J. Development of a self-report measure of capability wellbeing for adults: the ICECAP-A. Quality of Life Research. 2012; 21(1):167-176.

[6] Coast J, Flynn TN, Natarajan L, Sproston K, Lewis J, Louviere JJ, et al.: Valuing the ICECAP capability index for older people. Soc Sci Med. 2008; *67*(5):874-882.

[7] Hoefman RJ, van Exel J, Rose JM, Van De Wetering EJ, Brouwer WB. A discrete choice experiment to obtain a tariff for valuing informal care situations measured with the CarerQol instrument. Medical Decision Making. 2014 Jan;34(1):84-96.

[8] Makai P, Brouwer WB, Koopmanschap MA, Stolk EA, & Nieboer AP. Quality of life instruments for economic evaluations in health and social care for older people: a systematic review. Soc Sci Med. 2014;*102*:83-93.

[9] Knapp M, Mangalore R. The trouble with QALYs... . Epidemiol Psychiatr Soc. 2007;16(4):289-293.

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

[11] Brazier J & Tsuchiya A. Improving cross-sector comparisons: going beyond the health-related QALY. Applied health economics and health policy, 2015;*13*(6):557-565.

[12] Schawo S, Bouwmans M, Van der Schee E, Hendriks, V, Brouwer, W, Hakkaart, L. The search for relevant outcome measures for cost-utility analysis of systemic family interventions in adolescents with substance use disorder and delinquent behavior: A systematic literature review. Health Qual Life Outcomes. 2017;15(1):179.

[13] Goorden M, Schawo SJ, Bouwmans-Frijters CA, van der Schee E, Hendriks VM, & Hakkaart-van Roijen L. The cost-effectiveness of family/family-based therapy for treatment of externalizing disorders, substance use disorders and delinquency: a systematic review. *BMC psychiatry*. 2016;*16*(1):237.

[14] Brysbaert M. *Psychologie.* Gent: Academia Press; 2009.

[15] Cottrell D, Boston P. Practitioner review: The effectiveness of systemic family therapy for children and adolescents. J Child Psychol Psychiatry. 2002;43(5):573-586.

[16] Randall J, Cunningham PB. Multisystemic therapy: a treatment for violent substance-abusing and substance-dependent juvenile offenders. Addict Behav. 2003;28(9):1731-1739.

[17] Hendriks V, van der Schee E, Blanken P: Treatment of adolescents with a cannabis use disorder: Main findings of a randomized controlled trial comparing multidimensional family therapy and cognitive behavioral therapy in The Netherlands. Drug Alcohol Depend. 2011.

[18] Henggeler S W, Sheidow AJ. Empirically supported family-based treatments for conduct disorder and delinquency in adolescents. J Marital Fam Ther. 2002;38(1):30-58.

[19] French MT, Zavala SK, McCollister KE, Waldron HB, Turner CW, Ozechowski TJ. Cost-effectiveness analysis of four interventions for adolescents with a substance use disorder. J Subst Abuse Treat. 2008;34(3):272-281.

[20] Tanner-Smith EE, Wilson SJ, & Lipsey MW. The comparative effectiveness of outpatient treatment for adolescent substance abuse: A meta-analysis. Journal of substance abuse treatment, 2013;*44*(2):145-158.

[21] Darnell AJ, Schuler MS. Quasi-experimental study of functional family therapy effectiveness for juvenile justice aftercare in a racially and ethnically diverse community sample. Child Youth Serv Rev. 2015;50:75–82.

[22] Van der Stouwe T, Asscher JJ, Stams GJ, Deković M, van der Laan PH. The effectiveness of Multisystemic Therapy (MST): a meta-analysis. Clin Psychol Rev. 2014;34(6):468–81.

[23] Kaminer Y, Bukstein O, Tarter R E. The Teen-Addiction Severity Index: rationale and reliability. Int J Addict. 1991;26(2):219-226.

[24] Reckers-Droog V, Goorden M, Kaminer Y, van Domburgh, L, Brouwer, W & Hakkaart-van Roijen L. Presentation and Validation of the Abbreviated Self Completion Teen-Addiction Severity Index (ASC T-ASI): A Preference-Based Measure for Use in Health-Economic Evaluations. Submitted.

[25] de Bekker-Grob EW, Ryan M, Gerard K. Discrete choice experiments in health economics: a review of the literature. Health economics. 2012 Feb 1;21(2):145-72.

[26] Clark MD, Determann D, Petrou S, Moro D, de Bekker-Grob EW. Discrete choice experiments in health economics: a review of the literature. Pharmacoeconomics. 2014 Sep 1;32(9):883-902.

[27] de Bekker-Grob EW, Donkers B, Jonker MF, Stolk EA. Sample size requirements for discrete-choice experiments in healthcare: a practical guide. The Patient-Patient-Centered Outcomes Research. 2015 Oct 1;8(5):373-84.

[28] McFadden D. The measurement of urban travel demand. Journal of public economics. 1974 Nov 1;3(4):303-28.

[29] Hensher DA, Rose JM, Greene WH. Applied Choice Analysis. New York: Cambridge University Press; 2005.

[30] Vojacek O, & Pecakova I. Comparison of discrete choice models for economic environmental research. Prague Economic Papers, 2010;19(1):35-53.

[31] Centraal Bureau voor de Statistiek. Bevolking; kerncijfers. Retrieved January 1, 2013, from http://statline.cbs.nl/Statweb/?LA=nl

[32] Stolk E, Krabbe P, Busschbach J, Stolk EA, Krabbe P, Busschbach J. Using the Internet to collect EQ-5D norm scores: a valid alternative. In Proceedings of the Plenary Meeting of the EuroQoL Group, The Hague 2007; Sep 13.

[33] Sindelar JL, & Jofre-Bonet M. Creating an aggregate outcome index: cost-effectiveness analysis of substance abuse treatment. The Journal of Behavioral Health Services and Research, 2004; 31(3):229-241.

[34] Veldwijk, J, Lambooij MS, de Bekker-Grob EW, Smit HA, & De Wit GA. The effect of including an opt-out option in discrete choice experiments. PloS one, 2014;*9*(11), e111805.

[35] Versteegh MM, & Brouwer WBF. Patient and general public preferences for health states: a call to reconsider current guidelines. Social Science & Medicine. 2016;*165*:66-74.

Appendix A. English version of the ASC T-ASI instrument (B1 level) [22]

Please check the answer that currently fits you best:

1. Substance use

I have **no problem** with the use of alcohol, drugs or medicine ☐
I have a **slight problem** with the use of alcohol, drugs or medicine ☐
I have a **fairly large problem** with the use of alcohol, drugs or medicine ☐
I have a **large problem** with the use of alcohol, drugs or medicine ☐
I have a **very large problem** with the use of alcohol, drugs or medicine ☐

2. School

I have **no problem** with school ☐
I have a **slight problem** with school ☐
I have a **fairly large problem** with school ☐
I have a **large problem** with school ☐
I have a **very large problem** with school ☐

3. Work

I have **no problem** with work ☐
I have a **slight problem** with work ☐
I have a **fairly large problem** with work ☐
I have a **large problem** with work ☐
I have a **very large problem** with work ☐

4. Family

I have **no problem** with family ☐
I have a **slight problem** with family ☐
I have a **fairly large problem** with family ☐
I have a **large problem** with family ☐
I have a **very large problem** with family ☐

5. Social relationships

I have **no problem** with friends, acquaintances and others in my environment ☐
I have a **slight problem** with friends, acquaintances and others in my environment ☐
I have a **fairly large problem** with friends, acquaintances and others in my environment ☐

I have a **large problem** with friends, acquaintances and others in my environment ☐
I have a **very large problem** with friends, acquaintances and others in my environment ☐

6. Justice

I have **no problem** with the judicial authorities ☐
I have a **slight problem** with the judicial authorities ☐
I have a **fairly large problem** with the judicial authorities ☐
I have a **large problem** with the judicial authorities ☐
I have a **very large problem** with the judicial authorities ☐

7. Mental health

I have **no problem** with my mental health ☐
I have a **slight problem** with my mental health ☐
I have a **fairly large problem** with my mental health ☐
I have a **large problem** with my mental health ☐
I have **a very large problem** with my mental health ☐

Appendix B. Dutch version of the ASC T-ASI instrument (B1 level) [22]

Zet één kruisje bij het antwoord dat op dit moment het best bij jou past:

1. Middelengebruik

Ik heb **geen probleem** met het gebruik van alcohol, drugs of medicijnen ☐
Ik heb een **klein probleem** met het gebruik van alcohol, drugs of medicijnen ☐
Ik heb een **redelijk groot probleem** met het gebruik van alcohol, drugs of medicijnen ☐
Ik heb een **groot probleem** met het gebruik van alcohol, drugs of medicijnen ☐
Ik heb een **heel groot probleem** met het gebruik van alcohol, drugs of medicijnen ☐

2. School

Ik heb **geen probleem** met school ☐
Ik heb een **klein probleem** met school ☐
Ik heb een **redelijk groot probleem** met school ☐
Ik heb een **groot probleem** met school ☐
Ik heb een **heel groot probleem** met school ☐

3. Werk

Ik heb **geen probleem** met werk ☐
Ik heb een **klein probleem** met werk ☐
Ik heb een **redelijk groot probleem** met werk ☐
Ik heb een **groot probleem** met werk ☐
Ik heb een **heel groot probleem** met werk ☐

4. Familie

Ik heb **geen probleem** met familie ☐
Ik heb een **klein probleem** met familie ☐
Ik heb een **redelijk groot probleem** met familie ☐
Ik heb een **groot probleem** met familie ☐
Ik heb een **heel groot probleem** met familie ☐

5. Sociale relaties

Ik heb **geen probleem** met vrienden, bekenden en anderen in mijn omgeving ☐
Ik heb een **klein probleem** met vrienden, bekenden en anderen in mijn omgeving ☐
Ik heb een **redelijk probleem** met vrienden, bekenden en anderen in mijn omgeving ☐
Ik heb een **groot probleem** met vrienden, bekenden en anderen in mijn omgeving ☐

Ik heb een **heel groot probleem** met vrienden, bekenden en anderen in mijn omgeving ☐

6. Justitie

Ik heb **geen probleem** met justitie ☐
Ik heb een **klein probleem** met justitie ☐
Ik heb een **redelijk groot probleem** met justitie ☐
Ik heb een **groot probleem** met justitie ☐
Ik heb een **heel groot probleem** met justitie ☐

7. Geestelijke gezondheid

Ik heb **geen probleem** met mijn geestelijke gezondheid ☐
Ik heb een **klein probleem** met mijn geestelijke gezondheid ☐
Ik heb een **redelijk groot probleem** met mijn geestelijke gezondheid ☐
Ik heb een **groot probleem** met mijn geestelijke gezondheid ☐
Ik heb een **heel groot probleem** met mijn geestelijke gezondheid ☐

Appendix C. Example of a choice set (English translation)

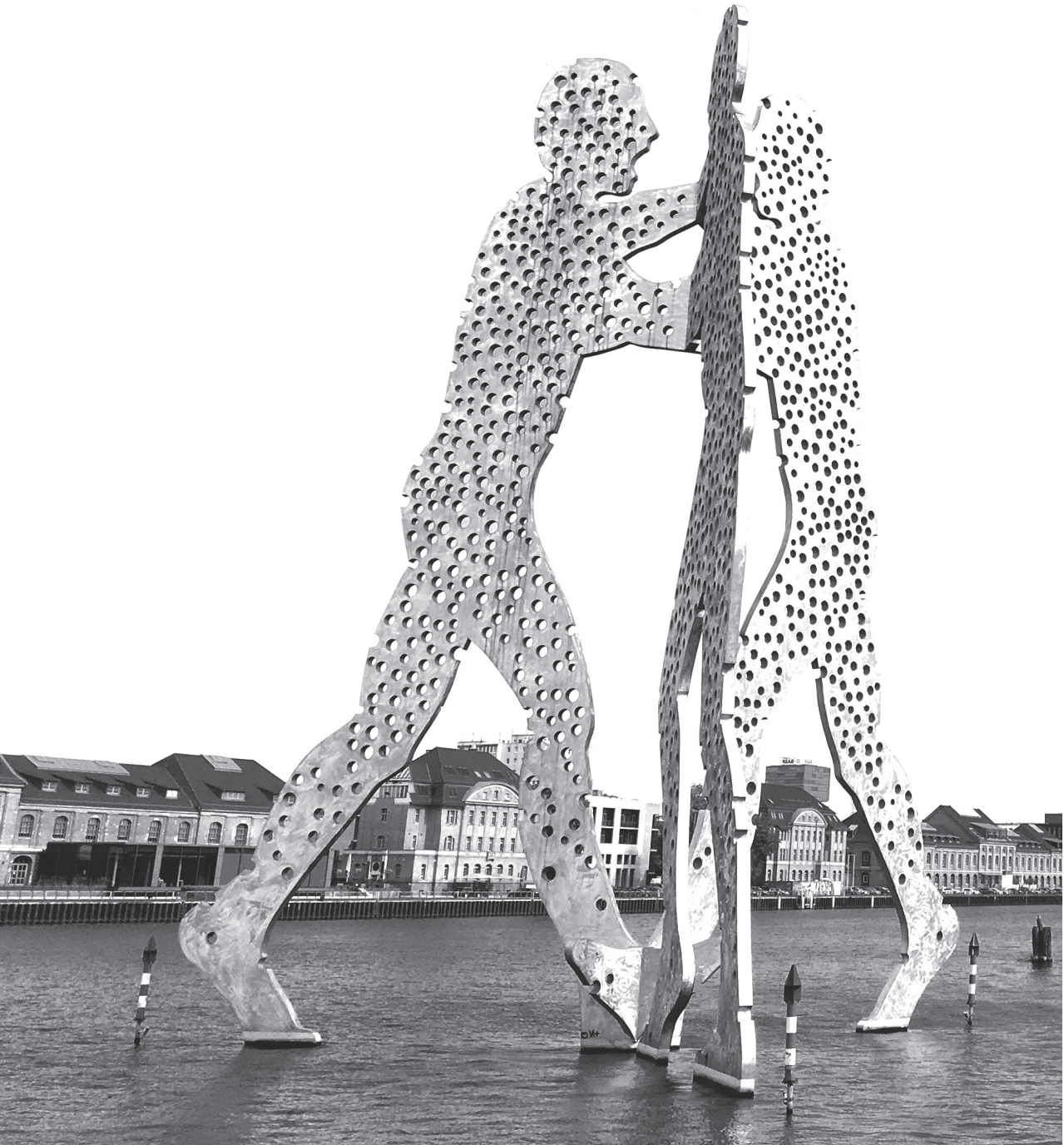
Which alternative do you consider best for the adolescent?

Below, two situations are presented in which a substance abusive and/or delinquent adolescent can end up: alternative A and B. Each alternative is further specified by the same seven problems (e.g. problems with school or within the family). However, the severity of these problems differs across the alternatives. The adolescent may experience **no problem, a slight, fairly large, large, or very large problem**.

What do we ask you to do? Think of a substance abusive and/or delinquent adolescent, whom you do not know. Then imagine that he experiences alternative A or B. Which alternative do you consider best for the adolescent, A or B? We will present you with 12 of these questions. Hence, you will have to choose 12 times between alternative A or B. Remember that all questions refer to the same adolescent in different situations. We are interested in your choice. There are no good or bad answers.
In case you would like more information on the type of problem you can move your cursor over the problem, e.g. 'school' to see more details.

Question 1. Which alternative do you consider best for the adolescent?

Alternative A	Alternative B
The adolescent has a **slight problem** with the use of alcohol, dugs or medicines	The adolescent has a **very large problem** with the use of alcohol, dugs or medicines
The adolescent has **a fairly large problem** with school	The adolescent has **a large problem** with school
The adolescent has **no problem** with work	The adolescent has a **slight problem** with work
The adolescent has a **very large problem** with his family	The adolescent has **no problem** with his family
The adolescent has **no problem** with friends, acquaintances and others in his surroundings	The adolescent has a **very large problem** with friends, acquaintances and others in his surroundings
The adolescent has a **very large problem** with justice	The adolescent has a **fairly large problem** with justice
The adolescent has a **slight problem** with his mental health	The adolescent has **a large problem** with his mental health



Chapter 9

Discussion

Introduction

The individual and societal impact of mental disorders is large. Effective treatment options can help to prevent mental disorders or reduce their impact. Yet, mental health treatments typically compete with other health care interventions for scarce financial resources. Hence, these interventions need to demonstrate their cost-effectiveness in order to ensure that they offer 'value for money' as compared to other treatment options. This is normally done through performing economic evaluations. As the results may inform funding and allocation decisions, it is crucial that such economic evaluations are executed well, fit to their precise purpose, and their results are carefully considered.

Measuring effects of mental health interventions as part of economic evaluations is not straightforward and much debated. The extent to which classical instruments to measure and value such effects (i.e., health economic outcome measures such as generic health-related quality of life questionnaires) adequately capture all relevant benefits of different mental health interventions has been questioned [1-3]. These concerns appear particularly relevant in the context of treatments for externalizing disorders in adolescents. These disorders can impact a broad array of life domains of the affected adolescents, and, also due to their life phase, strongly affect other people in their social environment. Patients may, for example, be acting out, pursue criminal activities, use drugs, or drop out of school, all of which result in costs and effects on and for the adolescents themselves as well as for the broader systems in which they operate (i.e., their family, school, neighborhood, etc.). Consequently, this may result in a high individual and societal burden within and beyond the health care sector. In assessing the impact of interventions aimed at treating adolescents with externalizing disorders, these broader costs and effects need to be recognized.

Treatments for externalizing disorders in adolescents include psychotherapeutic and psychosocial interventions as well as pharmacological treatments. As part of the non-pharmacological interventions, systemic family interventions have shown to be effective, yet are intensive and costly. Consequently, the question as to whether such interventions offer value for money is relevant. Economic evaluations attempting to assess this need to capture all relevant costs and benefits, including the broader impacts described above. The current methodology of economic evaluations does not appear to be fully adequate for doing so. This dissertation seeks to increase the awareness of this fact, as well as taking first steps in improving the methodology of economic evaluations. Our

aim was to explore various ways of expanding the scope of economic evaluations of interventions for externalizing disorders in adolescents. In particular, we investigated how cost-effectiveness analyses (CEA) of interventions for adolescents with externalizing behavioral disorders can be performed most meaningfully.

In the next section, I will highlight the results of the different chapters in relation to the specific research questions and the overall aim of this thesis. Thereafter, I will highlight important limitations and implications of the studies presented in this thesis.

Answering the research question

In **Chapter 2**, we first explored the use of existing health economic methodology with an extension to include broader societal outcomes. In doing so, we addressed the research question of *how to perform a CEA for pharmacological treatment of an externalizing disorder (ADHD), including consideration of relevant broader societal impacts while using existing health economic methodology.*

ADHD is common with a prevalence of 5% in children and adolescents [4] and is associated with broad societal and long-term effects [5-7]. Yet, the scope of the few available economic evaluations of treatments for ADHD is limited [8-10]. We performed an economic evaluation largely applying existing CEA methodology. To this end, we constructed a probabilistic Markov model to perform an illustrative calculation of cost-effectiveness of immediate release methylphenidate (IR) versus extended release methylphenidate (OROS). In contrast to most available economic evaluations, we included long-term effects and broader societal costs (e.g., costs of additional educational support and spillover effects in terms of caregivers' utility, medical costs and productivity losses). While the research presented in Chapter 2 highlighted that existing methodology can be used as a starting point for broader economic evaluations, in this particular case, available model input was limited, especially in terms of transition rates. Including broader elements of societal costs and benefits impacted the results of the economic evaluation, and arguably provided more relevant and disorder-specific estimates of cost-effectiveness. As such, our findings encourage adopting a broader perspective in CEA in the context of externalizing disorders.

In **Chapter 3**, we moved away from pharmacological treatments and towards systemic family interventions, the main focus of this dissertation. We performed a systematic literature review to investigate *the current knowledge regarding cost-effectiveness of systemic interventions for delinquency and substance use.*

Our final selection of papers included eleven studies. Because the results of those studies varied strongly – in terms of adopted perspective, which costs and effects were included, the treatment type and setting, and so forth – performing a meta-analysis was impossible. The quality of the included studies was found to be insufficient to draw firm conclusions about the cost-effectiveness of systemic interventions. As such, our exploration of the existing literature demonstrated that more research, of higher quality, is required. Not only do future studies need to clearly describe treatment types and use more sophisticated modeling approaches, they also need to take a broad (standardized) societal perspective in order to provide relevant estimates on the cost-effectiveness of systemic interventions.

In **Chapters 4 and 5,** we explored the possibility of adjusting the methodology of economic evaluations to better fit the characteristics and broad impact of systemic interventions for adolescents with problems of delinquency in particular. As a first step, we developed a simple, unidimensional overall outcome measure capturing broad outcomes and using this in the context of a CEA. We again constructed a probabilistic Markov model, this time to compare Functional Family Therapy (FFT) with treatment as usual. To capture the broad outcomes of systemic interventions in adolescents, we expressed the outcomes of the intervention in terms of the outcome 'Criminal Activity Free Years' (CAFY). This addressed the research question of *whether we can perform a CEA of a systemic intervention for delinquency in adolescents using Criminal Activity Free Years as an outcome measure.* The model included long-term effects and resulted in an estimate of incremental costs per CAFY. Using a probabilistic model and the CAFY outcome measure to assess cost-effectiveness of systemic interventions aimed to reduce delinquency was shown to be feasible. The presented model provided a framework to assess the cost-effectiveness of systemic interventions, while taking into account parameter uncertainty and long-term effectiveness. In **Chapter 5** we took this research one step further, using Value of Information analysis to address the question whether we can perform a Value of Information analysis based on the CEA using Criminal Activity Free Years as outcome measure to inform future research. We found that further research to eliminate parameter

uncertainty had a high value (i.e., €176 million) and that reducing uncertainty in some specific model parameters might be more valuable than in others. Such insights can steer future research, making it more efficient. Importantly, chapter 5 demonstrated that using a Value of Information framework to assess the value of conducting further research in the field of crime prevention, using a simple CAFY outcome measure, was indeed feasible. The results appeared relevant and well interpretable. However, notwithstanding the results from chapters 4 and 5, it should be noted that the CAFY (like, for example, event free life years) is a very simple and crude outcome measure, which is not preference-based (e.g., does not distinguish between types of delinquency). In that sense, though being more specifically directed at delinquency, it is a clear step back from generic health-related outcome measures. Moreover, its use will be hampered in practice by the fact that threshold values (i.e., what a CAFY would be worth to society) are not available. The use of an outcome measure like CAFY therefore results in problems common to CEA, using natural units as outcome measure (including problems of comparability of outcomes). This is not the case in cost-utility analyses (CUA) where standardized preference-based outcomes like QALYs are used.

Hence, the need for relevant, broad, yet more sophisticated, outcome measures, resembling some of the positive features of QALYs, is not reduced by having a CAFY measure. Therefore, in this thesis, attention was also paid to finding and developing preference-based outcome measures able to capture the relevant benefits of systemic mental health interventions. A logical starting point of course is to assert whether existing QALY measures indeed fail to sufficiently capture the goals and relevant effects of systemic mental health interventions. This can be done in a variety of ways. Here, we performed a qualitative study, presented in **Chapter 6**, in which we examined *which treatment effects should be captured in economic evaluations of systemic interventions in adolescents according to clinicians and whether the existing QALY measures capture these.*

Clinicians considered several EQ-5D dimensions relevant, in particular 'usual activities' and 'anxiety/ depression'. However, they also emphasized that the instrument lacked systemic dimensions such as family relations and relations with others, as well as addiction specific issues. These findings suggested that generic QALY measures like the EQ-5D instrument may not (directly) capture all relevant effects related to the here studied systemic interventions. This implies that economic evaluations using generic health-related quality of

life instruments as their outcome measure may omit relevant outcomes and, consequently, potentially lead to non-optimal policy decisions. More relevant outcome measures may be required, which could be either new measures or existing ones.

In **Chapter 7**, we therefore set out to investigate *which outcome measures are currently used to measure the effects of systemic interventions in clinical research and could be used in cost-utility analyses.* We performed a systematic literature review, which revealed a large variety of outcome measures currently used in effectiveness research of systemic interventions. However, only eight of the available instruments covered five or more life domains relevant in this context, and thus could be called comprehensive. None of these had preference-weights available. The T-ASI instrument was identified as promising in terms of comprehensiveness and potential to be transformed into a self-complete, preference based instrument for use in economic evaluations.

Based on the findings of **Chapter 7**, we investigated the question of *whether it is possible to obtain societal preference-weights for a comprehensive multidimensional outcome measure to be used in economic evaluations of systemic interventions targeted at adolescents with problems of substance use and delinquency.* Based on the findings presented in Chapter 7, an abbreviated self-complete version of the T-ASI (called ASC T-ASI) was developed [11]. **Chapter 8** describes how we derived societal preferences ('tariffs') for this instrument, which meanwhile has been validated with promising results [11]. We performed a discrete choice experiment (DCE) among a sample of the Dutch general population with the aim to obtain tariffs for the ASC T-ASI. Although respondents considered the choice task to be difficult, the results showed that the preference scores were logically ordered, with lower scores for worse states. All but one estimated coefficient were statistically significant. It turned out that problems concerning substance use, psychiatric status, and legal status were most influential, followed by family relations. School status, employment/support status, and peer/social relationships had less impact on overall scores. The obtained tariffs enable a preference-based assessment of broad effects of systemic family interventions for adolescents with problems of substance use and/or delinquency within economic evaluations when using the ASC T-ASI as outcome measure. However, given some methodological choices made in deriving these tariffs, more research is necessary to confirm the current tariffs in new and larger samples and to inform the interpretation of ASC T-ASI scores.

Concluding, we have explored various ways of expanding the scope of economic evaluations of interventions for externalizing disorders in adolescents throughout this dissertation. We started out with an exploration of classical CUA of pharmacological interventions extended with elements to account for broad societal costs. We then examined the available literature on cost-effectiveness and CUA of systemic family interventions. We observed a lack of state-of-the art CEAs as well as low quality of existing analyses of treatments for substance use and delinquency. In response, we performed a CEA of systemic family intervention and introduced a broad, yet basic outcome measure specific to delinquency. The evaluation was feasible, yet lacked the comparability allowed in cost utility analyses. Therefore, we explored next steps towards broader CUAs of systemic family interventions. To do this, we performed interviews among clinicians to establish whether the current QALY measure captures goals and relevant effects of systemic interventions and we investigated which outcome dimensions may be missing within the current methodology. The EQ-5D captured several relevant dimensions yet lacked others. In line with these findings we took steps towards developing a broader, more comprehensive outcome measure. We systematically searched for existing outcome measures with broad dimensions and suitable to be used to obtain societal preference weights. We identified a broad, established instrument, which is accepted in the field and covers the broader domains of interventions for externalizing disorders. Finally, we obtained societal preference weights for an abbreviated self-completion version of this instrument to facilitate its use in health economic evaluation.

Strengths and Limitations

Capturing the broad and (partly) non-medical effects of interventions for externalizing disorders in adolescents in economic evaluations has shown to be multi-faceted and challenging. Yet, information on the costs and effects of treatments for externalizing disorders is vital to the process of decision-making concerning resource allocation within the health care domain. We consider it a strength of this dissertation that various health economic techniques were explored and adapted in the context of the evaluation of systemic interventions of externalizing disorders. We built on existing data, added societal dimensions, examined a simple, yet broad measure of cost-effectiveness and explored possibilities for CUA by identifying a more comprehensive and meaningful measure and obtaining societal preference-weights for it. We have attempted to set first steps in including the broad effects of treatments for externalizing disorders in economic evaluations and have shown that the existing methodology,

together with extensions and a preference-based broad measure presented here, can be valuable in attaining this goal.

Several of the aspects discussed in this dissertation (e.g., broad effects on patient, third parties and society as a whole) may be common to more or even all mental disorders and may play a role in economic evaluations of these disorders as well. Family functioning, interaction with others and the ability to take social responsibilities, to participate in educational activities or the ability to engage in meaningful work will often be affected by a mental disorder. As such, these considerations should be factored in when determining the cost-effectiveness of interventions adopting a societal perspective. One could apply the classical health economic methodology using the QALY as outcome measure, possibly extended by measures to also capture broader effects or one could include these effects by choosing a broader outcome measure, which captures all relevant domains of functioning (like we presented here for delinquency and substance use). Thus, the findings from this dissertation may well be of value in thinking about ways to improve the health economic evaluation of these disorders in the future as well.

Yet, though our explorations resulted in valuable findings, many questions remain unanswered. We mention some, directly related to some of the studies performed in this thesis.

First of all, when investigating whether the QALY sufficiently captures relevant outcomes, we based our conclusions on literature and on findings from a qualitative study, whereas quantitative studies investigating the responsiveness of specific generic health-related quality of life instruments may have been valuable additions. Yet, those were beyond the scope of this dissertation.

Furthermore, and important also in terms of intended use of economic evaluations, though we and others [11] have taken several steps towards designing a comprehensive outcome measure for the economic evaluation of systemic interventions, we were not able to capture the specific outcomes of interventions without compromising on comparability of results. Within a clinical setting, CEA can be useful to compare similar interventions, yet it is limited in that outcomes are not comparable across interventions and across settings. Hence, CUA is more suitable for societal decision-making as it enables comparisons of different interventions to determine which intervention contributes most to health and

welfare compared to its costs. When using different 'utility measures', like the one investigated in this thesis, comparability is also sacrificed. In particular, results from the ASC T-ASI cannot be readily compared to QALY outcomes, as they measure different concepts and outcomes, and are expressed on different utility scales. (Note that this is also true for other broader outcome measures recently developed for economic evaluations, such as the ICECAP-O instrument or the ASCOT [12, 13]. This means that comparability is limited to other studies using the same or highly similar outcome measures. Hence, finding an adjusted outcome measure, like the ASC-T-ASI, which captures the broad and systemic effects of interventions means diminishing the comparability with outcomes of other economic evaluations. However, it also creates an opportunity to carefully and comprehensively capture the effects of the intervention, more in line with the intervention's goals and the results as viewed by patients and clinicians.

For the ASC T-ASI, an additional complication is that its tariffs are not expressed on an anchored scale like the QALY, so that 0 (1) simply means the worst (best) situation expressed with the instrument, making multiplications with duration less straightforward (also in terms of interpretation). In addition, no threshold values exist for the ASC T-ASI. Given the broad domains in the ASC T-ASI, specific attention needs to be paid to the possibility of double-counting items (on both the cost and the effect-side). Finally, even though first results on validity presented in [11] looked promising, to date the ASC T-ASI instrument has been validated only once, and so far no information is available on the sensitivity to change of the ASC T-ASI instrument. In that sense, the tariffs presented in Chapter 8, were developed soon after the first validation and do not imply that further validation is not encouraged. We also were not able to perform head-to-head comparisons of the ASC T-ASI instrument in a relevant sample with for instance conventional health related quality of life instruments like the EQ-5D, or more general preference based wellbeing measures, such as the ICECAP-A instrument [14].

Another important methodological challenge refers to the question of who should value the states described with a preference based outcome measure like the ASC T-ASI. In this dissertation, we opted for the common source of such valuations, i.e. the general population. While in valuing health states this already implies valuing hypothetical states (as many individuals will not be experiencing or even will ever have experienced the health state under valuation), here we asked respondents to choose what they believed to be best for the (hypothetical)

adolescent in that state. While one may feel this to be a logical choice, also for paternalistic reasons, other choices could have been made [15]. Future studies could investigate obtaining preferences of adolescent respondents (of the general population or even the target population) and see whether these preferences differ from the earlier obtained 'societal preferences'. In this way, we may obtain data from informants most closely related to the studied patients.

A further methodological limitation of the current thesis, also applicable to the evaluation of other mental disorders, is that we did not explicitly address the handling of common co-morbidities in performing health economic evaluations and interpreting their outcomes. In mental health, co-morbidities occur frequently and may affect the outcomes of a treatment. Treatment of one disorder may for instance positively or negatively affect the symptoms of a secondary disorder. When this is the case, observed treatment outcomes may be affected by the (possibly overlapping and interacting) symptoms of disorders. Furthermore, the presence of co-morbidity may complicate treatment and/or prolong the time to recovery. Consequently, taking a broad view on treatment outcomes will increase the chance of comprehensively capturing broad effects, yet interpreting the results is not always straight-forward in the presence of common co-morbidities and interdependencies, and may require close collaboration of clinicians, researchers and policy makers.

Some observations regarding the data used in the different studies also need to be made here. In the early chapters, we would have preferred (randomized control) trial data as input for the health economic models, yet unfortunately that data was not available and additional data collection exceeded the scope of this dissertation. Also, we would have preferred a larger number of participants for our interview study among clinicians, as a larger and more diverse group of respondents may have provided a more representative sample of systemic family therapists. The current sample may have lead to a relatively strong focus on drug use as the majority of respondents worked with patients with addiction problems. Additionally, it might have been interesting to have obtained information from other stakeholders, such as patients, to take an even broader perspective on the relevant outcomes of the interventions. Practical limitations and restrictions were the main reasons for not doing so. Furthermore, the discrete choice experiment performed within this dissertation was web-based and proved to be a difficult exercise for many respondents. The large number of participants of the discrete choice experiment who did not answer the control questions correctly and who

were therefore excluded from the analysis may have biased the results as the educational level of included respondents was shown to be higher than of the excluded individuals. Performing the experiment in a face-to-face setting may have improved the understanding of the respondents and may therefore have diminished drop-out and affected results.

Another noteworthy issue is that, in the current study, we obtained Dutch data only. Given this obvious limitation it may be interesting to investigate in future studies whether aspects like opinions of experts or societal weights for ASC T-ASI states differ in other countries. Differences may for instance occur, as substance use may be a more sensitive subject in some countries, responsibility of (young) individuals may be judged differently, importance of family relations may differ per country, etcetera. Consequently, this may affect what are considered important outcomes as well as the societal preference weights for these outcomes.

Recommendations for research and policy

Notwithstanding the limitations noted in the previous paragraph, we express several recommendations for future research. Policymakers increasingly face (difficult) decisions concerning the allocation of scarce health care resources. It is important that these decisions be informed in the best possible way. Health economic analyses in the context of externalizing disorders in adolescents have shown to be scarce and of insufficient quality, and common outcome measures seem too limited to capture relevant outcomes of the interventions. We explored different ways to expand the scope of economic evaluation of interventions for externalizing disorders in adolescents. Besides using and enhancing the common methodology, we also provided an alternative to the conventional QALY measure aimed at covering broader effects more in line with clinical goals.

Based on the findings of this dissertation, we suggest that economic evaluations of interventions for externalizing disorders in adolescents should, when performed according to conventional methodology, minimally include add-ons to account for broader elements of societal costs and benefits, without omissions or double-counting. These add-ons could be more limited additions to at least cover for instance the costs of additional educational support and spillover effects in terms of caregivers' utility, medical costs and productivity losses, as was done in Chapter 2. Ideally, we would recommend future trials to include broader outcome measures, such as the ASC-T-ASI, alongside common health-related quality of life measures. In this way, comparability with a wider range

of (medical) interventions would still be warranted based on the QALY outcome, yet one may also be able to use the more comprehensive ASC-T-ASI outcomes to make broader (systemic) effects visible and measurable, particularly useful in interpreting QALY results and when comparing with interventions in the same or a similar context. Directly investigating the needs and wishes of policy makers in this context, in terms of the type of information provided and the comparability of results, remains important as well. If comparability of outcomes of economic evaluations is compromised by using different utility measures in different contexts, this may pose less of a problem if appropriate 'thresholds' exist for these different outcome measures [16]. It would therefore be useful to define a threshold value for an improvement in functioning/wellbeing based on different preference-based outcome measures. This amount could be based on studies investigating how much people are willing to pay per point improvement on a relevant outcome measure. Such a threshold value is now increasingly investigated for QALY gains, but remains un(der)explored for other outcome measures, such as the ICECAP instruments, or the ASC T-ASI. If appropriate thresholds would exist, the issue of comparability of outcome measures is less problematic. We suggest future research on this topic such that outcomes can clearly be translated into policy advice.

Furthermore, we recommend future research to be directed at exploring different broad outcome measures, including the ASC T-ASI, further investigation their validity, sensitivity to change, and reliability, and confirming the current ASC T-ASI tariffs in larger samples. In addition, future research may be directed at obtaining preference scores in other countries to determine inter-country differences in societal preference scores, and obtaining preference scores in other samples (affected adolescents or a random sample of adolescents). Though this dissertation mainly focused on the effect side of the health economic evaluation of externalizing disorders in adolescents, various aspects connected to the cost side, such as inclusion of costs of crimes, incarceration, costs of lower level of education and earnings, etcetera, also warrant further investigation.

In addition, the research presented in this thesis may be valuable in the context of other areas and interventions in the field of mental health. The use of the ASC T-ASI or a similarly broad and preference-based measure in addition to the conventional QALY measure, in an attempt to better account for broad and diverse treatment effects, would be interesting in other areas as well. More

research on relevant outcomes and outcome measures in other areas of (mental) health care remains important.

General conclusion

When performing economic evaluations and attempting to assess the societal costs and benefits of interventions for externalizing disorders in adolescents, more suitable methodology is required. This dissertation has suggested the use of existing methodology, with (compared to common economic evaluations) additions and enhancements to better reflect the goals and outcomes of systemic interventions. In spite of our efforts, several questions remain unanswered. Yet, we have taken first steps in further raising awareness of this issue and in improving the methodology of economic evaluations in this field.

While working on this dissertation, I was particularly fascinated by the challenge of making outcomes of mental health inventions measureable, especially since such outcomes are often not visible from the outside, and not measurable by a single scan or by analyzing a single drop of blood. Nonetheless, the effects of interventions can make a huge difference for the wellbeing of a large number of individuals, together with the different systems they belong to. Furthermore, I was intrigued by the differences in viewpoint of clinicians and economists in this area, and felt an urge to bring these worlds closer together so as to improve their mutual understanding. I hope that this dissertation may contribute to the goal of making health economic evaluations of mental disorders, particularly with regard to externalizing disorders, more meaningful and, as such, serve as a starting point towards bridging the gap between health economic concepts and clinical practice in this field.

References

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

[2] Knapp M, Mangalore R. The trouble with QALYs. Epidemiol Psychiatr Soc 2007;16(4):289-293.

[3] Shah KK, Mulhern B, Longworth L, Janssen MF. Views of the UK General Public on Important Aspects of Health Not Captured by EQ- 5D. Patient 2017;13:1-9.

[4] Dalsgaard S, Østergaard SD, Leckman JF, Mortensen PB, Pedersen MG. Mortality in children, adolescents, and adults with attention deficit hyperactivity disorder: a nationwide cohort study. The Lancet. 2015 May 30;385(9983):2190-6.

[5] Tuithof M, Ten Have M, van Dorsselaer S, De Graaf R. Prevalentie, persistentie en gevolgen van ADHD in de Neder- landse volwassen bevolking. Tijdschrift voor Psychiatrie. 2014;56:10–9.

[6] Doshi JA, Hodgkins P, Kahle J, Sikirica V, Cangelosi MJ, Setyawan J, et al. Economic impact of childhood and adult at- tention-deficit/hyperactivity disorder in the United States. J Am Acad Child Adolesc Psychiatry. 2012;51(10):990e2–1002e2.

[7] Le HH, Hodgkins P, Postma MJ, Kahle J, Sikirica V, Setyawan J, et al. Economic impact of childhood/adolescent ADHD in a European setting: the Netherlands as a reference case. Eur Child Adolesc Psychiatry. 2013.

[8] Beecham J. Annual research review: Child and adolescent mental health interventions: a review of progress in economic studies across different disorders. J Child Psychol Psychiatry. 2014;55(6):714–32.

[9] Bernfort L, Nordfeldt S, Persson J. ADHD from a socio-eco- nomic perspective. Acta Paediatr. 2008;97(2):239–45.

[10] Wu EQ, Hodgkins P, Ben-Hamadi R, Setyawan J, Xie J, Sikirica V, et al. Cost effectiveness of pharmacotherapies for attention- deficit hyperactivity disorder: a systematic literature review. CNS Drugs. 2012;26(7):581–600.

[11] Reckers-Droog V, Goorden M, Kaminer Y, van Domburgh, L, Brouwer, W & Hakkaart-van Roijen L. Presentation and Validation of the Abbreviated Self Completion Teen-Addiction Severity Index (ASC T-ASI): A Preference-Based Measure for Use in Health-Economic Evaluations. Submitted.

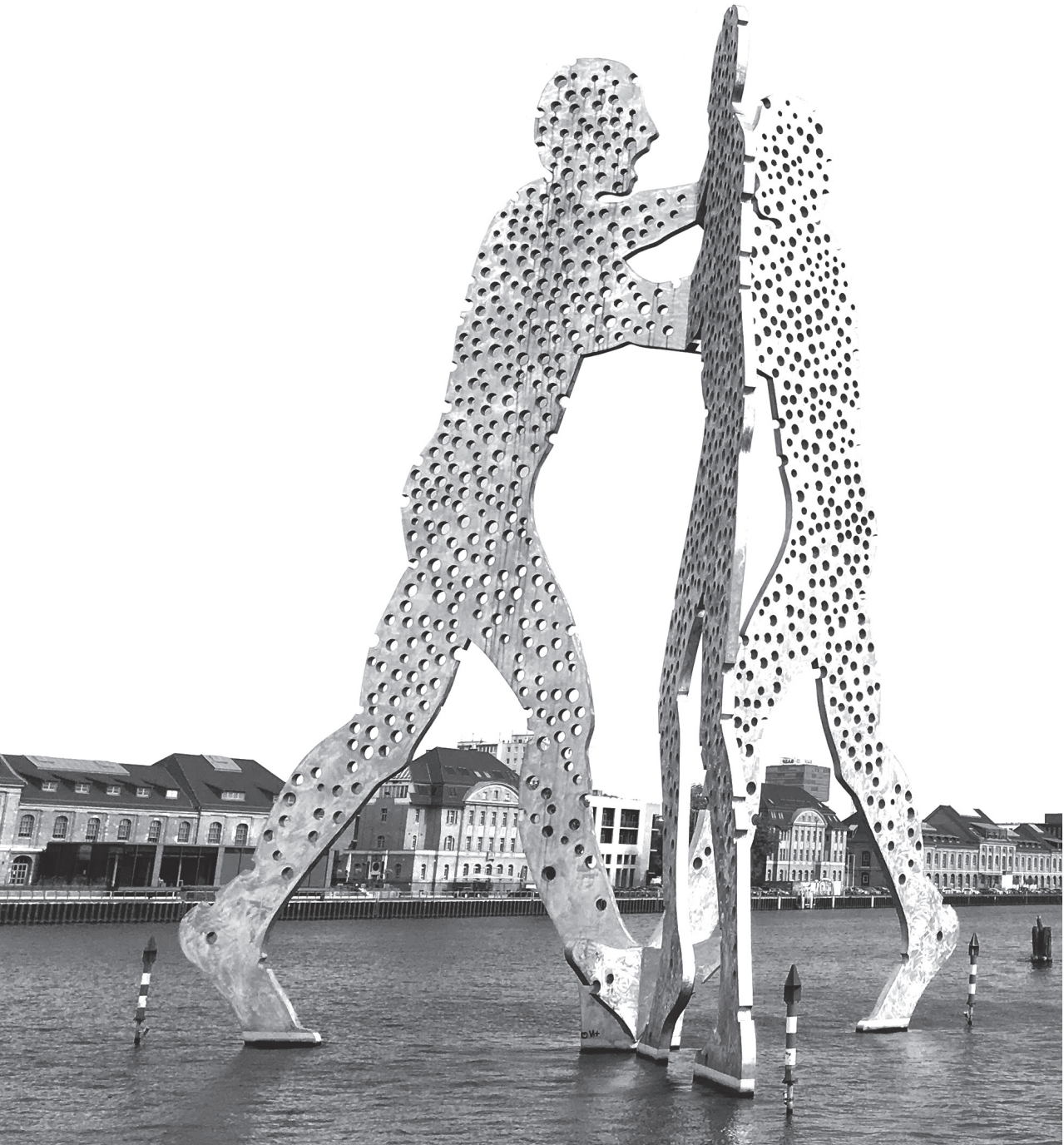
[12] Coast J, Peters TJ, Natarajan L, Sproston K, Flynn T. An assessment of the construct validity of the descriptive system for the ICECAP capability measure for older people. Quality of Life Research. 2008 Sep 1;17(7):967-76.

[13] Malley JN, Towers AM, Netten AP, Brazier JE, Forder JE, Flynn T. An assessment of the construct validity of the ASCOT measure of social care-related quality of life with older people. Health and Quality of life Outcomes. 2012 Dec;10(1):21.

[14] Al-Janabi H, Flynn TN, Coast J. Development of a self-report measure of capability wellbeing for adults: the ICECAP-A. Quality of Life Research. 2012 Feb 1;21(1):167-76.

[15] Versteegh MM, Brouwer WB. Patient and general public preferences for health states: a call to reconsider current guidelines. Social Science & Medicine. 2016 Sep 1;165:66-74.

[16] Brouwer W, van Baal P, van Exel J, Versteegh M. When is it too expensive? Cost-effectiveness thresholds and health care decision-making. 2019;175-180.



Appendices

Summary

Mental disorders are common. Effective treatments for mental disorders need to demonstrate their cost-effectiveness to compete for scarce, collectively financed health care resources. However, estimating the overall burden of these disorders and the effect of relevant treatments on patients, their environment and society as a whole are complex and challenging tasks. This particularly applies to externalizing disorders in adolescents where treatments often involve the systems around the patients and where a wide variety of societal costs and effects occur. Economic evaluations of such treatments do not yet seem to capture these broad and multifaceted costs and effects sufficiently. In this dissertation, we explored various ways to expand the scope of existing economic evaluations of interventions for externalizing disorders in adolescents to account for these broader costs and effects.

We started out by performing a 'conventional' cost-utility analysis (CUA) of pharmacological interventions extended with elements to account for broader societal costs. We then examined the available literature on cost-effectiveness analyses (CEA) and CUAs of systemic family interventions. Based on the literature, we observed a lack of state-of-the art economic evaluations as well as low quality of existing analyses of treatments for substance use and delinquency. Subsequently, we performed a CEA of a systemic family intervention and introduced a broad, yet very basic outcome measure specific to delinquency (i.e., criminal activity free years). Such evaluation proved to be feasible, yet lacked the comparability of outcomes facilitated by cost utility analyses. Therefore, we next explored steps towards broader CUAs of systemic family interventions. To do so, we first performed interviews among clinicians to establish whether the common QALY measure adequately captures all goals and relevant effects of systemic interventions. We also investigated which outcome dimensions may be missing within the current methodology (e.g. when using the EQ-5D instrument). Clinicians indicated that the EQ-5D captured several relevant dimensions yet lacked others. In line with these findings we took steps towards developing a broader, more comprehensive outcome measure. We systematically searched for existing outcome measures with broad dimensions and suitable to be used to obtain societal preference weights. We identified a broad, established instrument, which already is accepted in the field, and covers the broader domains that interventions for externalizing disorders aim to affect. Finally, we obtained societal preference weights for an abbreviated version of this instrument, suitable for self-completion, to facilitate its use in health economic evaluation.

We can conclude that when performing economic evaluations and attempting to assess the societal costs and benefits of interventions for externalizing disorders in adolescents, more suitable methodology is needed. In this dissertation, various health economic techniques were explored and adapted in this context. We have attempted to investigate how the broad effects of treatments for externalizing disorders can be included in economic evaluations and have shown that the existing methodology, together with extensions and a preference-based broad measure, can be valuable in attaining this goal. We have taken first steps in further raising awareness of this issue and in improving the methodology of economic evaluations in this field.

In spite of our efforts, several questions remain unanswered and some of the explored directions (such as using a new outcome measure) also raise new questions (e.g. regarding comparability of results with other interventions). Future research could further investigate these issues in order to improve the methodology of economic evaluations in the context of mental disorders in general and of externalizing disorders in adolescents in particular.

Samenvatting

Psychische aandoeningen hebben een hoge prevalentie. Behandelingen van zulke aandoeningen dienen aantoonbaar effectief en doelmatig te zijn om in aanmerking te komen voor schaarse, collectief gefinancierde middelen. Echter, het is een uitdagende taak om de kosten en effecten van de diverse behandelingen op de patiënt, zijn omgeving en de maatschappij als geheel accuraat in te schatten. Dit is in het bijzonder het geval voor externaliserende psychische aandoeningen bij adolescenten waarbij vaak 'het systeem rondom de patiënt' (zoals familie, buurt en school) bij de behandeling betrokken wordt en waarbij een breed scala aan maatschappelijke kosten en effecten ontstaan. Economische evaluaties van dergelijke behandelingen lijken op dit moment deze brede en uiteenlopende maatschappelijke kosten en effecten veelal buiten beschouwing te laten. In dit proefschrift is getracht bij te dragen aan de verbreding van de reikwijdte van economische evaluaties van behandelingen voor externaliserende problematiek bij adolescenten, teneinde het meewegen van bredere maatschappelijke kosten en opbrengsten te bevorderen.

Allereerst is een conventionele kosten-utiliteiten analyse (KUA) van farmacotherapeutische interventies uitgevoerd, waaraan enkele elementen van bredere maatschappelijke kosten werden toegevoegd in de analyse. Vervolgens hebben wij de bestaande literatuur over kosteneffectiviteitsanalyses (KEAs) en kostenutiliteitsanalyses (KUAs) van systeeminterventies bij externaliserende problematiek bij adolescenten, met name middelengebruik en delinquentie, gereviewd. Hierbij kwam naar voren dat er een gebrek is aan economische evaluaties op dit terrein en dat de schaarse gepubliceerde evaluaties vaak te wensen over laten in termen van kwaliteit. Daarna hebben wij een KEA van systeeminterventies uitgevoerd en hierbij een brede, zeer eenvoudige uitkomstmaat, specifiek voor delinquentie geïntroduceerd (namelijk criminaliteit-vrije-jaren). De evaluatie met behulp van deze uitkomstmaat bleek uitvoerbaar, echter werd noodzakelijkerwijs hierdoor de vergelijkbaarheid van de uitkomsten met andere evaluaties verminderd en maakt de uitkomstmaat geen onderscheid tussen verschillende vormen van criminaliteit en is deze beperkt qua inhoud.

Derhalve zijn in het proefschrift aansluitend eerste stappen gezet om bredere KUAs van systeeminterventies mogelijk te maken. Allereerst hebben wij daartoe interviews met clinici gehouden om informatie in te winnen over de vermeende geschiktheid van een veelgebruikte uitkomstmaat die het mogelijk maakt QALYs te berekenen, de EQ-5D, om doelen en relevante effecten van systeeminterventies

te meten. Verder hebben wij de vraag voorgelegd welke dimensies volgens hen ontbraken in een conventionele economische evaluatie, waarin het EQ-5D instrument als uitkomstmaat wordt gebruikt. Clinici gaven aan dat de EQ-5D meerdere belangrijke dimensies omvat, echter dat er ook een aantal belangrijke dimensies ontbraken. In het verlengde van deze bevindingen hebben wij stappen gezet richting de ontwikkeling van een bredere, meer omvattende uitkomstmaat. We hebben systematisch gezocht naar bestaande instrumenten met brede dimensies die geschikt zouden zijn voor het verkrijgen van maatschappelijke waarderingsscores. We hebben een breed, bestaand instrument geïdentificeerd, dat reeds geaccepteerd is in het veld en dat de bredere domeinen afdekt die van belang zijn voor interventies voor externaliserende problematiek. Tenslotte hebben we maatschappelijke gewichten voor een verkorte versie van het instrument (waarbij de patiënt het instrument zelf invult) verkregen om het gebruik in economische evaluaties beter mogelijk te maken.

We kunnen concluderen dat beter passende methodologie nodig is voor het uitvoeren van economische evaluaties van interventies gericht op het behandelen van externaliserende problematiek, indien het doel is om alle maatschappelijke kosten en effecten mee te wegen. Binnen dit proefschrift hebben we in deze context diverse economische technieken geëxploreerd en aanpassingen voorgesteld. We hebben gepoogd te onderzoeken hoe de brede effecten van interventies voor externaliserende problematiek meegenomen kunnen worden binnen economische evaluaties en laten zien dat de bestaande methodologie, met verschillende aanvullingen en een maatschappelijk gewaardeerde uitkomstmaat, waardevol kunnen zijn om dit doel te bereiken. Daarmee zijn eerste stappen gezet om het bewustzijn met betrekking tot deze onderwerpen te vergroten en om de methodologie in dit veld te verbeteren.

Ondanks deze resultaten blijven veel vragen onbeantwoord en ook roepen sommige van de genoemde mogelijkheden (zoals het gebruik van een nieuwe uitkomstmaat) nieuwe vragen op (bijvoorbeeld met betrekking tot de vergelijkbaarheid van de uitkomsten met andere interventies). Toekomstig onderzoek zou dieper op deze vraagstukken kunnen ingaan teneinde de methodologie van economische evaluaties in de context van geestelijke gezondheid nog verder te verbeteren.

List of publications

Scientific publications

Schawo S, Hoefman R, Reckers-Droog VT, van de Wetering L, Kaminer Y, Brouwer WBF, Hakkaart-van Roijen L. Obtaining preference scores for an abbreviated self-completion version of the Teen-Addiction Severity Index (ASC T-ASI) to value therapy outcomes of systemic family interventions: a discrete choice experiment. *submitted*.

Schawo S, Carlier I, Van Hemert A, De Beurs E. Measuring treatment outcome in patients with anxiety disorders – a comparison of the responsivity of generic and disorder-specific instruments. *Journal of anxiety disorders*. 2019;*64*: 55-63.

Schawo S, Brouwer W, Hakkaart L. Clinicians' views on therapeutic outcomes of systemic interventions – which are relevant outcomes and do common health economic evaluations capture them? The Journal of Mental Health Policy and Economics. 2017;20(3): 131-136.

Schawo S, Bouwmans C, van der Schee E, Hendriiks V, Brouwer W, Hakkaart L. The search for relevant outcome measures for cost-utility analysis of systemic interventions in adolescents with substance use disorder and delinquent behavior: A systematic literature review. Health and Quality of Life Outcomes. 2017;15(1):179.

Goorden M*, Schawo S*, van der Schee E, Hendriiks V, Hakkaart L. The cost-effectiveness of family interventions on substance abuse and delinquency: A Systematic review. BMC Psychiatry. 2016;16(1):237.

van Eeren H, Schawo S, Hakkaart L, Busschbach J. Value of information analysis applied to systemic interventions aimed to reduce juvenile delinquency. PlosOne. 2015;10(7).

van der Kolk A, Bouwmans C, Schawo S, Buitelaar KJ, van Agthoven M, Hakkaart-van Roijen L. Association between societal costs and treatment response in children and adolescents with ADHD and their parents. A cross-sectional study in the Netherlands. SpringerPlus. 2015;4:224.

Schawo S, van der Kolk A, Bouwmans C, Annemans L, Postma M, Buitelaar J, van Agthoven M, Hakkaart-van Roijen L. Probabilistic Markov Model Estimating

Cost Effectiveness of Methylphenidate Osmotic-Release Oral System Versus Immediate-Release Methylphenidate in Children and Adolescents: Which Information is Needed? PharmacoEconomics. 2015;33(5):489-509.

van der Kolk A, Bouwmans C, Schawo S, Buitelaar J, van Agthoven M, Hakkaart-van Roijen L. Association between Quality of Life and Treatment Response in Children with Attention Deficit Hyperactivity Disorder and their Parents. The Journal of Mental Health Policy and Economics. 2014;17(3):119-29.

Bouwmans C, van der Kolk A, Oppe M, Schawo S, Stolk E, van Agthoven M, Buitelaar J, Hakkaart-van Roijen L. Validity and responsiveness of the EQ-5D and the KIDSCREEN-10 in children with ADHD. The European Journal of Health Economics. 2013;15(9): 967-77.

Schawo S, van Eeren H, Soeteman D, van der Veldt MC, Noom M, Brouwer W, Van Busschbach J, Hakkaart L. Framework for Modelling the Cost-effectiveness of Systemic Interventions Aimed to Reduce Youth Delinquency, The Journal of Mental Health Policy and Economics. 2012;15:187-196.

Hakkaart-van Rooijen L, Tan SS, Goossens L, Schawo S, Brouwer W, Rutten-van Molken M, Rutten F. Doelmatigheid in praktijkrichtlijnen voor medicijnen: resultaten van een 'quick scan'. TSK. 2010;88 (4):175-181.

Other publications

Bouwmans C, Schawo S, Jansen D, Vermeulen K, Reijneveld S, Hakkaart-van Roijen, L. Vragenlijst & Handleiding Intensieve jeugdzorg. Zorggebruik en productieverlies; 2012.

PhD portfolio Saskia Schawo

PhD training

2008	Socio economic evaluation of medicines - Quality of Life (University of York, UK)
2008	Advanced Modelling Methods for Economic Evaluation (University of Glasgow, UK)
2009	Discrete Choice Models - Theory and Applications to Environment, Landscape, Transportation and Marketing (University of Bologna, Italy)
2009	Principles of Epidemiologic Data-analysis (NIHES Research Training in Medicine and Health Sciences, Rotterdam)
2009/2010	Writing Academic English for iBMG staff (Academic Language Centre, Leiden University)
2010	'Klaar in vier jaar' (Brigitte Hertz, Rotterdam)
2011/2012	Discrete Choice Models – course for iBMG staff (University of Sydney, hosted at Erasmus University Rotterdam)
2013	Psychiatric Epidemiology (NIHES Research Training in Medicine and Health Sciences, Rotterdam)
2014	'Inleiding in de psychologie' (Open University Heerlen)
2014	'Biologische Grondslagen: Neuropsychologie en Psychofarmacologie' (Open University Heerlen)
2015	Pre-master 'Klinische Psychologie' (Leiden University)
2016	MSc. Clinical Psychology, cum laude (Leiden University)
2018	'Basiscursus CGT 100 uur' (Fortagroep, Rotterdam)
2018	'Basiscursus Interpersoonlijke Therapie' (Parnassia Academy, The Hague)
2019	'Basiscursus Schematherapie' (Parnassia Academy, The Hague)

Teaching

2010/2011	'MS-Excel vaardigheden', practicum, bachelor program Health Sciences, Institute of Health Policy and Management, Erasmus University Rotterdam
2013	'Methoden en Technieken van Sociaal Wetenschappelijk Onderzoek 1', practicum, bachelor program Health Sciences, Institute of Health Policy and Management, Erasmus University Rotterdam
2013/2014	'Financieel Management', practicum, master program Health Care Management, Institute of Health Policy and Management, Erasmus University Rotterdam

Conferences

Podium presentations

2009 9th Workshop on Costs and Assessment in Psychiatry, Venice, Italy

2010 8th European Conference on Health Economics, Helsinki, Finland 2009

2013 11th Workshop on Costs and Assessment in Psychiatry, Venice, Italy

Poster presentations

2008 ISPOR 11th Annual European Congress, Athens, Greece

2010 11th Biennial International EUSARF conference, Groningen

Other meetings and workshops

2010, 2011 Lowlands Health Economic Study Group, Egmond aan Zee

Curriculum vitae

Saskia Schawo was born on July 21st, 1980 in Bonn, Germany. She attended high school in Germany, during which time she spent a year studying abroad.

After graduating from high school in 1999, Saskia came to the Netherlands to study Economics at Maastricht University. During her studies, she worked as an intern at Allianz Insurance in both Munich and Rotterdam. In 2004, she received her master's degree in Economics and started to work at Allianz Nederland as a financial controller. During the next four years, she performed solvency and risk calculations and became interested in mathematical modeling.

Between 2008 and 2014, Saskia shifted her attention toward the intersection between health and economics. As a researcher and Ph.D. candidate at the Erasmus School of Health Policy and Management, her research focused on health economic evaluations of mental health interventions. In particular, Saskia became interested in possibilities for economic evaluation in the field of mental health that could capture clinical treatment goals while making use of existing health economic methodology. Saskia also taught scientific methods and Financial Management to students as part of her Ph.D. training.

In 2014, with an interest in mental disorders peaked by her research and a growing interest in clinical practice, Saskia decided to study Clinical Psychology at Leiden University. She received her Master's degree, cum laude, in 2016. and has since been working as a psychologist at Parnassia Groep in The Hague and in Zoetermeer. She has been treating adults with ADHD, anxiety and mood disorders.

Alongside her studies and work, Saskia has volunteered at Debora (a center for cancer patients and their families) in Delft and at Rivierduinen, a mental health institution in Leiden.

Dankwoord

Dit proefschrift is ontstaan in een bewogen tijd. En er zijn veelvuldig momenten geweest waarop de nadelen van het afronden groter leken dan de voordelen. Dit boekje was dan ook niet tot stand gekomen zonder de waardevolle steun, hulp en persisterende aanmoediging van een aantal bijzondere mensen.

Werner en Leona, bedankt voor jullie waardevolle en kritische commentaar, jullie gedrevenheid, beschikbaarheid en betrokkenheid tijdens het hele traject. Bedankt ook voor de gezelligheid en de muzikale onderbouwing van onze promotie overleggen.

Ik wil de promotiecommissie bedanken voor het beoordelen van mijn proefschrift, het waardevolle commentaar en de gedachtewisseling tijdens de verdediging.

Dit onderzoek was niet tot stand gekomen zonder de welwillende en gemotiveerde therapeuten en patiënten van de Bascule, de Parnassia Groep en de Viersprong die de tijd hebben genomen om de deels zeer uitgebreide vragenlijsten te beantwoorden en die hiermee een waardevolle bijdrage hebben geleverd aan dit onderzoek. Hetzelfde geldt voor de respondenten van de online panels, de deelnemers van het expert panel en de medewerkers van Survey Sampling International. Bedankt voor jullie inzet en prettige samenwerking.

Tijdens het onderzoek heb ik met veel collega's van binnen en buiten de Erasmus Universiteit samengewerkt. Ik wil mijn medeauteurs en oud-collega's bedanken voor de fijne samenwerking, de ontspannen sfeer, de openheid en gezelligheid (met en zonder kroket) en de mooie momenten op de diverse congressen in het buitenland. In het bijzonder: lieve Hester, Jan, Clazien, Maartje, Renske, Henk, Caroline, Margreet, Sofie, Annemieke, Laurens, Pepijn, Saskia de G., Pieter, Siok Swan, Kim, Tim, Martine, Liesbet, Lucas, Arthur en Melinde: jullie hebben mijn werkdagen gezellig en bijzonder gemaakt!

Daarnaast wil ik mijn (oud-) collega's bedanken voor de gezelligheid en steun de afgelopen jaren. Lieve Han Lin en Wouter, bedankt voor de fijne band die we met elkaar hebben en de gezellige lunches, wel of niet bij V&D. Lieve oud-collega's van PsyQ Den Haag, hartelijk dank voor jullie vertrouwen in mij, voor de kansen binnen een nieuw vakgebied, de gezelligheid, jullie vriendschap en bedankt voor alles dat ik van jullie heb mogen leren. Lieve collega's van PG

Zoetermeer, ik geniet er elke dag opnieuw van om met jullie samen te mogen werken. Jullie hebben mij me meteen welkom en thuis laten voelen. Dank voor jullie betrokkenheid, jullie gezelligheid, interesse en vriendschap.

Lieve vrienden, in het bijzonder Uli, Pascal, Hannah, Sacha, Walter, Menno, Lisette, Jur, Anne B., Karel, Rob, Cis, Stef, Danker, Vanessa W. en Suzanne, bedankt voor het fijne contact, de verbondenheid, jullie interesse, luisterende oor, jullie aanmoedigingen, steun en vertrouwen. Ik ben erg blij en dankbaar dat jullie er zijn.

Lieve Gees en schildervriendinnen van de Keet, ik geniet van de fijne schilderavonden met jullie; dank voor alle gezelligheid, creativiteit, verbondenheid, inspiratie en vrijheid om mezelf te kunnen zijn. Dank ook voor jullie support m.b.t. de laatste loodjes en jullie kritische kijk op de portretten.

Lieve Anne v.D. en Janneke, ik ben niet alleen erg dankbaar voor jullie vriendschap en de dierbare en mooie momenten die we samen beleven maar ik vind het ook heel fijn en bijzonder dat jullie vandaag mijn paranimfen willen zijn. Dat betekent veel voor mij.

Lieve Vanessa v.N., als er een persoon is die gezorgd heeft dat ik dit boekje af heb kunnen maken, dan ben jij het. Je bent niet alleen een lieve en slimme vriendin maar je kreeg het ook voor elkaar om mij telkens weer te motiveren om door te gaan – met succes. Dank daarvoor!

Lieve Annemarie, dank je wel voor je speelse en zachte aanwezigheid, je luisterende oor, je rust, kalmte, inspiratie en steun op elk moment. Ik kan moeilijk verwoorden wat dit voor mij betekent.

Lieve Carolien, dank voor je vriendschap en de leuke, avontuurlijke en gezellige uitstapjes samen in binnen- en buitenland en bedankt dat je je kennis en kunde van Photoshop hebt ingezet om de voorkant van het boekje samen verder uit te werken.

Lieve Julie, dank voor je vriendschap, de mooie tijd op Buitenkunst en in Denemarken, en bijzondere dank voor de correctie van mijn Engelse teksten.

Lieve 'schoonfamilie', we hebben de afgelopen jaren samen veel meegemaakt en ik ben blij en dankbaar dat er als vanzelfsprekend een mooie band blijft, dat we elkaar weten te vinden, dat we bijzondere momenten met elkaar delen en vieren.

Lieve mama en papa, ik wil jullie bedanken voor de kansen die jullie voor mij hebben gecreëerd, de vrijheid die ik heb gekregen om het buitenland in te trekken en mij hier te ontwikkelen. Ik vind het fijn dat ik de rust van de "Heimat", de zachte heuvels, de Duitse taal, de krakelingen en het Duitse brood, de appelbomen en het dennenbos op kan zoeken wanneer ik daar behoefte aan heb.

Lieve Helm, we hebben samen ervaren hoe het is als er geen tijd meer is, hoe je alles zou willen geven om nog maar één of twee of drie QALY's erbij te krijgen. We hebben keihard gevochten, maar ook gelachen. We hebben genoten van wat er wel was. Ik heb die weg voortgezet en je hebt me geholpen om koers te zetten richting wat ik waardevol en belangrijk vind om mijn tijd aan te besteden. Kon ik de dag vandaag maar met je delen.