The impact and treatment of developmental stuttering

Caroline de Sonneville

The Impact and Treatment of Developmental Stuttering

The Impact and Treatment of Developmental Stuttering

De impact en behandeling van ontwikkelingsstotteren

Proefschrift

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

Prof.dr. H.A.P. Pols

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

donderdag 29 oktober 2015 om 13.30 uur

door
Caroline de Sonneville-Koedoot
geboren te Rotterdam

Erasmus University Rotterdam

PROMOTIECOMMISSIE

Promotoren
Prof.dr. C.A. Uyl-de Groot
Prof.dr. R.J. Baatenburg de Jong

Overige leden
Prof.dr. J.J.V. Busschbach
Prof.dr. F. Verhulst
Prof.dr. B.A.M. Maassen

Copromotoren
Dr. E.A. Stolk
Dr. M.C. Franken

CONTENTS

1

Introduction

BACKGROUND AND MOTIVATION

In the last decade, the field of developmental stuttering is gradually evolving into an evidence based practice (EBP) discipline. EBP is a derivative of the concept of evidence based medicine (EBM) and is defined as "the integration of the best research evidence with clinical expertise and patient values, aiming to provide optimal clinical care" [1]. In order to practice EBP, clinicians have to identify and apply high-quality scientific evidence in the process of clinical decision making. Despite an increasing research effort within paramedical disciplines, about a decade ago the evidence base underlying speech and language pathology was still scarce [2]. Most studies included small sample sizes and only applied condition-specific outcome measures. Decision making by speech-language pathologists (SLPs) was primarily opinion- and experience-based. The growing emphasis on empirically supported treatments affected directions of research in the field developmental stuttering, as evidenced by a special edition of the *Journal of Fluency Disorders* addressing EBP in 2003 [3-5] and publication of the results of the first phase III randomized trial into stuttering treatment in children in 2005 [6].

The need for rigorous outcome evaluations in the field of developmental stuttering is embedded in a changing health care environment. In particular, rising medical costs have placed an increasing emphasis on *value for money* in reimbursement decisions and, consequently, on the empirical validation of treatments. Although speech and language pathology in the Netherlands had not yet been affected by health care budget restraints at the time this thesis was initiated, a future demand for more evidence on effectiveness and cost-effectiveness was anticipated.

Thus, there is an increased demand for more studies in the field of developmental stuttering in order to enhance clinical care and strengthen the position of SLPs in reimbursement decision-making. It is against this background that this thesis found its origin. Before describing the topics of this thesis, the context in which this work was conducted will be discussed. Specifically, I will focus on the current state of knowledge on the disorder and treatment of developmental stuttering, and the need for incorporating broad based health outcome measures as well as costs in the evaluation of stuttering.

DEVELOPMENTAL STUTTERING

Stuttering is a developmental disorder characterized by an abnormally high frequency of disruptions in the flow of speech. These disfluencies can take the form of repetitions of sounds and syllables ("W-W-W-Where are you go-go-going?"), prolongations ("SSSSave

me a seat."), or blockages (no sound) [7,8]. People who stutter may use excessive physical and/or mental effort in order to speak fluently. Speech may therefore be associated with unusual facial and body movements. To the speaker, stuttering can involve far more than just the observable behavior [9]. People who stutter may for instance experience feelings of shame, embarrassment, and anxiety due to their stuttering. Stuttering is most likely to start in the preschool years. The peak onset is around the age of 2.6-3.6 years [10,11]. About 5 to 11% of children in the age of 3 to 6 years start to stutter [11,12]. Since about 175.000 children are born in the Netherlands every year, this implies that each year 8750 to 19000 Dutch children are affected.

The precise cause of stuttering is yet not fully understood, although research from the last two decades has resulted in a great leap forward into the knowledge about the etiology of stuttering. Whereas sixty years ago stuttering was perceived to be caused by parents labelling normal disfluencies in the child's speech as stuttering (Wendell Johnson's "diagnosogenic theory"), nowadays stuttering is commonly viewed as a multifactorial disorder. It is most probably the result of a complex interaction between genetic susceptibility to the disorder, neurological development, environmental factors and child developmental factors [8,13], although it is not clear how each of these factors precisely contribute to the onset of stuttering. Family studies [14,15], twin studies [16,17], adoption studies [18] and, more recently, genetic linkage studies [19-22] have shown evidence for a strong genetic basis underlying the stuttering disorder. Evidence for a neurological component includes studies showing neuroanatomical differences [23,24] and differences in in auditory processing of speech and language by children who stutter (CWS) compared to children who do not stutter (CWNS) [25,26], as well as differences in neural pathways involved in speech fluency [27]. Examples of environmental factors that may precipitate stuttering in children prone for the disorder are high parental expectations and a hectic, fast-talking home. Lastly, research findings suggest that the speech language system of CWS is more fragile and susceptible to disruption than that of CWNS. Studies have pointed to difficulties in lexical encoding [28], the influence of sentence length and syntactic complexity on stuttering [29,30], and discrepancies among language skills in stuttering children [31,32]. Studies into the relation between phonology and stuttering have provided mixed results [33]. Reilly et al. [11] showed that stuttering children had *better* language skills than their non-stuttering peers, suggesting that stuttering might be a "by-product" of rapid language development in the preschool years.

Most children outgrow the disorder of stuttering before the age of 10 years [12,34,35], resulting in a prevalence rate of about 1% in the adult population [7]. The natural recovery rate (thus, without formal intervention) about 5 years after onset is estimated to be 70 to 85% [12,34,36,37]. Girls, children who start to stutter before the age of 3 years and children

with a family history of recovery from stuttering have a higher chance to recover naturally [38]. Recovery may be initiated close to the onset of the disorder, however, in most cases full recovery takes several years. Although several predictive factors for recovery are known, it remains unpredictable whether an individual child will recover spontaneously or not. For those who do not recover, the risk of negative effects hampering psychosocial development lies ahead. For instance, school aged children who stutter have shown to be more vulnerable to bullying [39]. Furthermore, stuttering has been linked to higher levels of social anxiety in adults [40]. Persistent stuttering can impact on several domains of functioning and quality of life, like social and emotional functioning [41] and employment opportunities [42,43]. The risk of these negative long-term effects warrant treatment close to the onset of stuttering. Moreover, treatment in the preschool years is associated with a higher chance on successful outcome, presumably so because neural plasticity decreases with age. Delaying treatment beyond 15 months post onset is known to reduce the chance of full recovery by about 25% [44]. Hence, early intervention is generally recommended for children who stutter.

TREATMENT OF DEVELOPMENTAL STUTTERING IN PRESCHOOL CHILDREN

In the Netherlands, parents of children suspected of stuttering will usually consult a general practitioner or an SLP. Treatment is, however, not necessary for every child. Many children go through a period of speech disfluencies while developing their speech and language skills. The SLP's assessment will include evaluation of the frequency, type, duration and severity of disfluencies, and observation of excessive physical effort, signs of frustration or other emotional reactions by the child. These signs, together with the presence or absence of risk factors for persistency and parental concerns, help the clinician in the process of identifying children in need for treatment.

Two basic approaches to treatment for preschool children who stutter can be discriminated. *Indirect* approaches focus on manipulating child related and environmental factors assumed to influence the child's speech fluency, while *direct* approaches directly target the child's speech fluency [44-46]. An example of the first approach is treatment based on the Demands and Capacities Model (DCM) [47], which has been the standard treatment for preschool stuttering children in the Netherlands since the late eighties. An example of the latter approach is behavioral treatment according to the Lidcombe Program for early intervention (LP) [48]. The LP was introduced in the Netherlands in the year 2000, based on promising initial results in Australia [49-51]. In the next section indirect and direct treatment for preschool stuttering children will be discussed. DCM based treatment and the LP will receive particular attention.

Indirect treatment: Treatment based on the Demands and Capacities Model

Indirect approaches to stuttering in children are based on the theoretical notion that stuttering is a multifactorial disorder [52-54] with physical, linguistic, psychological, and/or environmental factors influencing the onset and development of stuttering. Specifically, most indirect treatments find their origin in the Demands and Capacities model, which will be discussed next.

Demands and Capacities Model (DCM) based treatment is premised on the assumption that stuttering develops when a child lacks the capacities to speak as fluently as the environment demands. Therefore, treatment aims to achieve a favorable balance between environmental demands and demands by the child him- or herself, and the child's capacities for speaking fluently [47,55]. Demands and capacities can be motoric, linguistic, emotional or cognitive of nature. For instance, parents are trained to slow down their habitual speech rate (motoric demand), reduce their number of questions (linguistic demand), respond to the child's specific temperament (emotional demand), or ask questions of conceptual complexity that is age-appropriate (cognitive demand). Examples of the child's capacities that are addressed are improving the child's speech motor movements (motoric capacity), training word finding capacity (linguistic capacity), promoting his self-esteem (emotional capacity) or teaching the child the concept of turn-taking and rules for conversation (cognitive capacity).

Prior to therapy, demands and capacities are assessed by way of speech- and language assessments and a videotape of parent-child interaction. In a first parent session, parents are informed on stuttering and their concerns are explored. Parents are also introduced to the concept of "parent-child special time". That is, they are required to spend 15 minutes a day (for minimal 5 days a week) giving the child their undivided attention and practicing homework assignments. Treatment generally starts with lowering demands through counseling and training of the parents, the child himself and significant others in the child's environment. Subsequently, explicit training of the capacities of the child may be introduced. If lowering the demands and promoting the capacities should fail to resolve the stuttering problem to a satisfactory extent, speech fluency may be worked on by modelling slower, more relaxed, smoother speech. The parents and the child initially attend the clinic once a week for a one-hour session. After four sessions with parent and child, as a rule a parent session will take place at which the child is not present. The intensity of therapy is gradually reduced when the child shows acceptable speech, parents master implementing a fluency enhancing environment and know what to do if a relapse occurs. The mean number of sessions is estimated to be 12, however with a high variability [47]. In a preliminary study by Franken, Kielstra, and Boelens [56], the mean number of treatment sessions of DCM based treatment of one hour each in the first three months was 11, and

three-fourths of children were still on treatment at the end of the three months. In the rest of this introduction I will refer to DCM based treatment as RESTART-DCM treatment; the DCM based approach that was evaluated in the RESTART-trial of which the results are presented in this thesis.

Examples of other published indirect treatment approaches are Parent Child Interaction Therapy (PCIT) [57,58] and Family Focused Treatment [59]. They share several components with RESTART-DCM treatment. All treatments begin with a comprehensive assessment, based on the multifactorial model of stuttering. This assessment helps the therapist and parents to identify the factors that are supporting or impacting the child's fluency, forming the basis for setting up treatment goals. Furthermore, parents play an important role throughout the whole course of therapy. Environmental or parental aspects are typically addressed first, before the child-focused aspects of treatment (i.e., the child's capacities). Lastly, the treatments are generally designed to be flexible in order to adapt to the specific strengths and needs of the child and family. There are also differences between these indirect treatments. For instance, PCIT and Family-focused treatment place more emphasis on the "demand" side. Within PCIT, this mainly consists of improving the quality of the parent-child relationship and changing parent-child interaction patterns. The central part of Family-focused treatment is educating parents on stuttering and teaching them to implement facilitating communication modifications in their interactions with their child skills. PCIT and Family-focused treatment also deploy a more restricted structure than RESTART-DCM treatment. PCIT is delivered as six once weekly clinic sessions, followed by a 6-week home consolidation period [60,61]. Family-Focused treatment usually consists of six to eight 45-minute appointments and is divided into three sections: (a) education and counselling of parents, (b) communication modification training of parents, and (c) review and reassessment [59].

Direct treatment: the Lidcombe Program for early intervention

Direct stuttering treatment approaches are not explicitly based on a theory of the onset and development of stuttering, but hold the notion that manipulating the child's speech production will increase fluent speech and decrease stuttered speech. Direct treatment for preschool children who stutter mainly include response-contingent therapies. The best-developed and most extensively researched form of response-contingent treatment for children is the Lidcombe Program, which will be discussed next.

The Lidcombe Program (LP) [48] is a behavioral treatment for preschool children who stutter, based on operant methods. In the LP, parents are taught by the SLP to deliver verbal contingencies during conversations with their child when the child is speaking mostly stutter-free. Contingencies for stutter-free speech are acknowledgment or praise

of fluency (e.g., "That was smooth," "Good talking!") or requesting self-evaluation (e.g., "Were there any bumpy words?"). Contingencies for stuttered speech include acknowledgment of stuttering (e.g., "That was a bit bumpy.") or requesting self-correction (e.g., "Can you say that again?"). There should be far more contingencies for stutter-free speech than for stuttered speech. Initially, contingencies are given in daily, so-called *structured conversations* of about 15 minutes. When the parent has become proficient in delivering contingencies and the stuttering of the child during the day is mild, the parent starts delivering contingencies during daily conversations. The treatment consists of two stages. During stage 1, the parent conducts the treatment every day and the parent and child attend the speech clinic once a week. This continues until stuttering either disappears or reaches an extremely low level. During stage 2, the use of parental feedback as well as the number of clinic visits is gradually reduced, provided that fluency is maintained. The LP is individualized for every child and family (e.g., according to the child's age or stuttering severity).

Examples of other direct treatment programs for preschoolers who stutter, although less frequently applied, are the Fluency Development System for Young Children [62], Extented Length of Utterance (ELU) [63], and Gradual Increase in Length and Complexity of Utterance (GILCU) [64]. Although the Fluency Development System for Young Children is mainly a direct approach (i.e., teaching children to use smooth and relaxed speech instead of bumpy speech), it also comprises indirect elements, like parent training. The ELU and GILCU are fluency shaping programs for older children. In an adapted form, they can be applied in preschool children. They are, as the LP, based on operant conditioning, but explicitly structures treatment by gradually increasing the length and complexity of utterances.

All Dutch SLPs are trained in DCM based treatment in their regular education program. Therefore, in Dutch primary care most children who stutter are treated according to a DCM approach. However, the methods used in the various education programs differ, presumably causing non-uniformity in clinical practice. A smaller number of SLPs has also followed Lidcombe Program training. Thus, the choice for treatment according to a DCM approach or the LP depends for a great part on the clinician a child presents to. It was only recently that the first Dutch guideline on stuttering was published [65]. This guideline can help in clinical decision making; however, most of the recommendations in this guideline are based on limited evidence.

PLUGGING GAPS IN THE EVIDENCE BASE UNDERLYING TREATMENT FOR PRESCHOOL CHILDREN WHO STUTTER

Current evidence on the effectiveness of treatments

Randomized Clinical Trials (RCTs) are considered the gold standard for proving clinical research evidence. A recent review by Nye and colleagues [66] identified seven studies into stuttering treatment outcome in preschool children that applied an RCT design. The mean quality of these studies was judged to be moderate. Moreover, six of these studies dealt with the efficacy of the LP [6,56,67-70]. Except for the study by Jones et al. [6], these studies had a short follow-up (4 to 16 weeks) and small sample sizes (12 to 45 stuttering children). The follow-up in Jones et al. [6] lasted nine months and was adequately powered (n=110), however, only 54 children (49%) were included [6]. The results of the studies by Jones et al. [6], Harris et al. [67], and Lattermann et al. [69] showed that the LP leads to a larger and faster reduction in stuttering frequency than natural recovery. However, there is a lack of replications by independent researchers as well as prospective studies with a follow-up that lasts long enough for children to complete treatment. Moreover, the long-term results obtained with the LP have not been compared to that of clinically relevant alternative interventions. Franken et al. [56] compared the outcomes of the LP and DCM based treatment in 23 children in a 12-week period. Stuttering frequency decreased for both treatments after 12 weeks and there was no between-group difference. This study was the pilot-study of the RESTART-trial, for which the results are presented in this thesis. Compared to the evidence base for the LP, the scientific evidence underlying indirect treatment is marginal. Besides the preliminary study by Franken et al. [56], no studies employing an RCT design have been performed. Two single subject studies exploring the efficacy of PCIT provide some empirical support for the PCIT [60,61]. A preliminary study examining the outcomes of Family-Focused Treatment in 17 preschool children demonstrated reduced frequency of stuttering-like disfluencies at the end of treatment and at long-term follow-up [59]. Thus, with the limited data available at present, the LP offers the best evidence-based intervention for preschool CWS [66]. Based on the evidence presented above, the Dutch guideline for stuttering (Richtlijn stotteren) [65] recommends that both the LP and DCM based treatment can be considered in selecting a treatment approach.

The need for incorporating broad outcome measures

Since the speech disruptions and concomitant stuttering behaviors are the central feature of the stuttering disorder, it is not surprising that treatment evaluation has traditionally focused on these observable characteristics. Moreover, the selection of treatment outcomes generally is closely related to the goal of treatment. For children who stutter, the primary goal of treatment approaches is to reduce or eliminate the stuttering behavior.

1

Therefore, it is understandable that measures such as percentage of syllables stuttered (%SS) or a severity judgment by the clinician or parent have been conventionally used in the evaluation of therapy for preschool stuttering children. On the contrary, the goal of most treatments for older children and adults who stutter is not only to reduce the observable stuttering behavior, but also to diminish any negative psychosocial impact. There are several instruments to evaluate non-speech related aspects, like the Speech Situation Checklist [71] to evaluate speech anxiety, the Erickson S24 [72,73] to evaluate communication attitude, and the Subjective Screening of Stuttering [74] to evaluate the speakers' social, emotional, cognitive, or other reactions to their stuttering. These instruments have, however, all been criticized for their weak psychometric properties [75]. Moreover, their scope is limited to one or a few aspects of the broader impact of the stuttering disorder.

Since the year 2000, the potential of broad-based outcome parameters like quality of life and well-being have been acknowledged in the field of stuttering research [41,75-78]. This aligns with a general shift in health care outcome research in the last decade from symptoms as the primary outcomes toward direct measurement of the impact of the disorder on the patient's daily life. This shift is motivated by acknowledgment of the complex relationship between biological aspects, symptoms and the impact on daily living [79,80] and studies showing that changes in physical endpoints are often only partly related to changes in patient evaluated health status [81]. Furthermore, studies in the field of mental health have shown that it is the patients' subjective well-being, rather than the objective medical condition, that determines their treatment-seeking behavior, their compliance and their evaluation of treatment [82]. Lastly, in policy decisions on treatment reimbursement, the burden of a disease has gained importance in relation to considerations on acceptable cost-effectiveness ratios. That is, the higher the burden of disease, the more willing society is to accept unfavorable ratios of costs to effects. For these reasons, it has become recognized that patient reported outcomes such as quality of life should be incorporated in treatment evaluation.

The growing awareness of the importance to incorporate broader outcome measures in the evaluation of therapies has affected the last-decades research topics in the field of stuttering. An increasing number of studies into the impact of stuttering by quality of life or well-being instruments have been undertaken [41,42,83-85], in which both generic and disease-specific instruments have been applied. An example of a disease-specific instrument developed in the last decade is the OASES [86].This questionnaire evaluates the physical, social and emotional aspects of stuttering from the perspective of the individual who stutter. The OASES for adults (OASES-A) has become a widely-applied diagnostic clinical tool to assess the broad spectrum of the stuttering disorder. Although disease-specific instruments like the OASES may provide valuable insights into the effect

of stuttering on multiple aspects of daily living, their usefulness is limited in that they focus on the domains of quality of life that are most likely to be influenced by the disease. Instead, generic instruments incorporate general health-related physical, social and emotional aspects of quality of life. These generic health-related quality of life (HRQOL) instruments are therefore capable to measure the burden of a disease, regardless of the underlying diagnosis. In this way, the broad impact of a disorder could be measured and the disease burden of stuttering could be compared with that of other conditions. This could help the disorder of stuttering in gaining a firm foothold in the current dynamic health care climate. The number of studies applying generic HRQOL instruments to assess the impact of persistent stuttering was scarce at the onset of this thesis. In the pediatric field of stuttering, and speech- and language problems in general, the application of this kind of measures has also not gained much attention [87]. Furthermore, generic HRQOL instruments have not yet been applied in the evaluation of stuttering treatment, while this kind of instruments provides essential outcome information necessary in cost-effectiveness analysis [88]. The next section will explain how the measurement of HRQOL in relation to costs could be helpful in the evaluation of stuttering treatment.

The relevance of economic evaluations

Rapidly increasing health care expenditures forces policy makers to make explicit decisions on the expenses of the health care budget. The most efficient allocation of finite health care resources requires not only information on the outcomes of interventions, but also consideration of related costs. In other words, interventions need to be judged on their relative *value for money*. The method of economic evaluations has become an increasingly important tool for priority settings in health care. In an economic evaluation the costs and outcomes of alternative health care interventions are compared [88]. Contrary to the medical field, in which economic evaluations have become quite well established, in the field of speech and language pathology economic evaluations are relatively uncommon. This may partly be related to the disease burden associated with most of the speech and language disorders expected to be relatively small, compared to medical disorders like cancer or heart failure. Moreover, the costs associated with speech and language disorders are a relatively small part of the health care budget.[1] It is probably because of these reasons that stuttering treatment (and speech and language therapy in general) so far has escaped from budget restrictions set by Dutch policy makers. In light of the pressing health care budget, however, evidence on the relative costs and effects of stuttering treatments are crucial to substantiate decisions regarding choices on and reimbursement of treatments.

In an economic evaluation, two or more alternative health care interventions are compared.

---

1 Based on an estimation of at least 2500 children in the Netherlands who undergo treatment each year, the total annual health care costs are estimated to be at least E1,5 million.

Usually, one intervention is a newly introduced intervention, and the other(s) consist(s) of usual care. The most common forms of an economic evaluation are the cost-effectiveness analysis (CEA) and a subtype of this analysis: the cost-utility analysis (CUA). In a CEA, effects are expressed in clinically relevant outcome measures related to the intervention. In the case of stuttering treatment, possible outcomes include (the reduction of) percentage syllables stuttered (%SS) or stuttering severity. In a CUA, the health outcome is the QALY, which stands for Quality Adjusted Life Years. The QALY combines length of life and quality of life in a single outcome measure. Where efficacy and effectiveness studies investigate the benefit of a treatment for the individual, cost-effectiveness studies investigate the benefit for society. CEA's and CUA's are therefore preferably undertaken from a societal perspective, that is; including all relevant costs and effects, no matter whom they are related to. Costs and effects of alternative health care interventions are compared through the calculation of an incremental cost-effectiveness ratio (ICER). This ratio expresses the additional costs of an intervention per additional unit of health gain on a clinical outcome (in case of a CEA) or per QALY gained (in case of a CUA). Since the QALY outcome in a CUA allows for comparison across health states, this type of economic evaluation is preferred for health care decision making. If the ICER of a new health intervention compared to usual care is below a decision maker's willingness-to-pay threshold, the new intervention is considered to be cost-effective.

OUTLINE OF THIS THESIS

The previous introduction showed the gap in the scientific evidence base underlying the treatment of developmental stuttering. In order to provide treatment based on the highest standards of care and to meet demands of policy makers, data on the impact of stuttering and the costs and effects of treatment is urgently needed. The main focus of this thesis is therefore on the impact and treatment of developmental stuttering.

The first three chapters address the impact of stuttering. **Chapter 2** explores the impact of the disorder in early childhood, by evaluating the HRQOL of preschool children who stutter. For this study, baseline HRQOL data of CWS participating in the RESTART-study and data of a Dutch population of CWNS were compared. In addition, the relation between stuttering severity and HRQOL was evaluated. **Chapter 3** explores the impact of persistent stuttering on daily life, by assessing the generic HRQOL of a convenience sample of adults who stutter. Evaluation of the HRQOL associated with persistent stuttering in adults will shed light on the disease burden we would like to prevent by treating stuttering in the preschool years. Since stuttering in adulthood results from a complex interaction between the core stuttering behaviors and emotional, cognitive and behavioral learning processes, it can

be assumed that psychological factors play an important role in the relationship between stuttering severity and HRQOL in adults. Therefore, the study described in chapter 3 applied a comprehensive approach to investigate the relationship between stuttering and HRQOL and incorporated the role of coping style. The Dutch OASES for adults (OASES-A-D) was used to assess the experience of the stuttering disorder from the perspective of the adult who stutter. **Chapter 4** reports on the translation process and evaluates the psychometric properties of the OASES-A-D.

The last three chapters of this thesis describe the results of the RESTART-trial (the *R*otterdam *E*valuation study of *S*tuttering *T*herapy in preschool children- *A R*andomized *T*rial). In this 18-months multicenter randomized trial the Lidcombe Program (LP) was compared with RESTART-DCM based treatment in 199 preschool CWS. Besides the comparison of the clinical effectiveness, an economic evaluation was conducted. **Chapter 5** presents the results on the relative effectiveness of both therapies. The primary clinical outcome for this study was the percentage of children recovered from stuttering after 18 months. **Chapter 6** reports on the cost-effectiveness and cost-utility of the LP in comparison with RESTART-DCM based treatment. Although the pragmatic design of the RESTART-trial enhanced the external validity, perceptions and beliefs of clinicians with regard to the treatments undoubtedly influence if and how they will incorporate the trial results. **Chapter 7** therefore describes the results of a focus group meeting, in which participating SLPs discussed their ideas and experiences with the LP and RESTART-DCM treatment in the context of the RESTART-trial.

Chapter 8 provides an overall discussion of the results of the previous chapters, including limitations, implications for clinicians and policy makers, and suggestions for future research.

To note, the chapters of this thesis are based on research articles published in or submitted to scientific peer reviewed journals. Therefore, the chapters can be read independently and some overlap exists between chapters.

2

Health-related quality of life of preschool children who stutter

With Elly Stolk
Hein Raat
Clazien Bouwmans
Marie-Christine Franken

ABSTRACT

Purpose The purpose of this study is to compare the health-related quality of life (HRQOL) of preschool children who stutter (CWS) and a reference population of children who do not stutter, and to evaluate the association between stuttering severity and HRQOL.

Methods Baseline data were used from 197 children participating in a multicenter Randomized Clinical Trial in the Netherlands. Information on stuttering severity and time since onset (TSO) of stuttering was obtained from the baseline evaluation by speech-language pathologists. Stuttering severity was measured using the SSI-3. HRQOL was assessed using proxy versions of two Child Health Questionnaires (ITQOL-97 and CHQ-PF28), the Health Utility Index 3 (HUI3) and the EuroQoL EQ-VAS (EQ-VAS).

Results While the outcomes on the EQ-VAS and the HUI3 showed that the HRQOL of CWS is slightly poorer than that of the Dutch reference population, results on the different dimensions of the CHQ-instruments did not reveal any difference in scores between stuttering children and reference groups. Within the group of CWS, two ITQOL-97 and four CHQ-PF28 scales showed statistically different scores for children in different SSI stuttering severity or TSO categories. However, the effect sizes showed that these differences were so small that they could be considered negligible.

Conclusion The results of this study do not reveal a diminished HRQOL for preschool CWS. Future research should include a larger cohort of children with severe stuttering, study the longitudinal course of HRQOL and incorporate additional parameters such as the characteristics of the child and his environment.

2

INTRODUCTION

Stuttering frequency and stuttering severity have traditionally been the primary outcome measures for childhood stuttering. However, these measures provide little information about the potentially broader impact of this disorder on daily life. Although health-related quality of life (HRQOL) measures have become an essential outcome in pediatric disorders [89,90], limited research has been conducted into the impact of speech and language disorders in general and stuttering in particular [87].

So far, research on the burden of stuttering has mainly focused on the impact of persistent stuttering. Several studies report a diminished functioning and/or (health-related) quality of life for stuttering adolescents and adults (e.g., [41,42,84,85,91]). For school aged children who stutter (CWS), the impact of stuttering on HRQOL has been less frequently studied. In fact, we are only aware of the study by Gooding and Davis [92], in which the generic HRQOL of 33 stuttering children and matched controls aged 8-16 years was assessed by the Child Health Questionnaire (CHQ-PF50). Their results did not reveal any significant group differences on general health, but significantly more behavioral problems were reported in stuttering children compared to matched controls.

Other studies applied disease (or disorder-) specific instruments to investigate the impact of stuttering on the daily life of school aged children. Chun, Mendes, Yaruss, and Quesal [93] showed a moderate negative impact of stuttering as measured by a draft version of the Overall Assessment of the Speaker's Experience of Stuttering for school aged children (OASES-S) in seven children aged 7 to 12 years, with a tendency toward a positive correlation between stuttering severity and the impact of stuttering on quality of life. Cook, Donlan, and Howell [94] also used a disorder-specific instrument (Fragebogen Zum Sprechen; FZS speech questionnaire) to assess the psychosocial impact of stuttering in 54 children aged 9 to 20 years, and found that higher stuttering severity was correlated with a greater psychosocial impact. Furthermore, Kawai, Healey, Nagasawa, and Vanryckeghem [95] and Vanryckeghem, Brutten, and Hernandez [96] showed a negative speech-associated attitude for school aged CWS compared to non-stuttering peers, and Blood and Blood [39] reported increased vulnerability to bullying for school aged CWS. A negative speech-related attitude and negative social reactions increase the child's risk for developing social anxiety [97]. Indeed, increased levels of social anxiety have been reported in older CWS, albeit not consistently [97]. A social anxiety disorder is known to hamper normal social development and functioning and is likely to result in a diminished quality of life [40].

In the age group where stuttering starts to develop, i.e., the preschool years, Reilly et al. [11] were the first to apply a validated descriptive HRQOL instrument (PedsQL). The

results of their population-based study indicated that stuttering in young children was not associated with a diminished quality of life. Surprisingly, for CWS aged 4 years, a better social and preschool functioning was reported compared with non-stuttering children. In addition, a study into social anxiety in preschool CWS by van der Merwe, Robb, Lewis, and Ormond [98] reported no significant differences between stuttering preschoolers and matched controls. However, other studies suggest that stuttering in children under the age of six may be related to psychosocial impairment. CWS as young as 3 or 4 years of age seem to be aware of their speech problems and, on average, evaluate their speech more negatively than their non-stuttering peers [96]. Awareness of stuttering in young children increases with time since onset and age, and stuttering could provoke frustration and facilitate emotional and behavioral reactions even in early childhood (e.g., [99]). Furthermore, differences in temperament characteristics and emotional behavior between preschool stuttering and non-stuttering children were found by various researchers (for an overview, see [100,101]). These characteristics (such as exhibiting poorer adaptability skills and more negative emotions) might amplify the psychosocial effects of stuttering, although speculations as to this connection are highly premature. Recently, a study by Kefalianos, Onslow, Ukoumunne, Block, and Reilly [102] suggested that preschool CWS do not have innately different temperaments from control children.

Negative reactions by listeners may also give rise to psychosocial consequences of stuttering. Negative evaluation of perceptually salient stuttering by preschool peers has been reported by Ambrose and Yairi [103] and Ezrati-Vinacour, Platzky, and Yairi [104]. While these studies were based on research using puppet-play, one study based on real-life interactions with peers also reported that preschool classmates reacted negatively to moments of severe stuttering [105]. Besides, the authors of the latter study observed that the stuttering preschool children had some difficulties in expressing themselves in social interactions, for instance in trying to take the lead in play or contributing to problem-solving activities. In a further study by Langevin, Packman, and Onslow [106], which focused on the parental evaluation of the impact of stuttering on the lives of their preschool stuttering children (N=77), the majority of parents reported that stuttering had a negative impact on their child's life. Both in children and parents, emotional consequences were reported. For example, children became frustrated because of their stuttering or had a low self-esteem, and stuttering affected the children's general mood. However, only 8% of the parents perceived that stuttering affected their child's quality of life [106].

From this overview it can be concluded that, although several studies suggest impairment in the psychosocial domains of HRQOL in preschool CWS, so far few attempts have been made to perform a comprehensive study to measure this impact with validated generic HRQOL instruments. These instruments allow comparison of the HRQOL between

preschool children with and without stuttering and thereby provide a basis to evaluate possible consequences of stuttering across multiple clinically relevant domains. These insights will ultimately enable therapists to optimize treatment goals. Therefore, the aim of the current study is to explore the HRQOL of preschool CWS using validated generic instruments.

METHODS

Participants

The study population of CWS consisted of preschool children participating in a multicenter Randomized Clinical Trial in the Netherlands named *RESTART* (the Rotterdam Evaluation study of Stuttering Therapy in preschool children- a Randomized Trial) which compares two alternative therapies for preschool CWS. All children presented to one of the 20 participating speech clinics for stuttering treatment. For the current study, baseline data obtained prior to randomization were used. Ethical approval for the RESTART-study was gained from the Medical Ethical Committee of the Erasmus Medical Center in Rotterdam (MEC-2006-349). Baseline data were collected between September 2007 and June 2010. All children met the following inclusion criteria: (1) between the ages of 3.0 and 6.3 years[1]; (2) stuttering confirmed by a rating of stuttering severity on an 8-point scale of at least 2 ("mild") [34] by the parent (3) as well as by the clinician; (4) children stuttered at least 3% of syllables; and (5) time since onset (TSO) of stuttering at least six months. Children were excluded if an emotional, behavioral, learning or neurological disorder had been diagnosed, or if there was a lack of proficiency in Dutch for children and parents. Written informed consent was obtained from all parents.

In total, 199 CWS were included in the RESTART-study. Two children were excluded from the current study because no baseline HRQOL data were available. Thus, this study was based on data of 197 CWS: 146 children aged 3 to 4 years (66% boys) and 51 children aged 5 to 6 years (80% boys).

HRQOL reference values

To compare HRQOL values of CWS with reference values, available data from two representative community samples of Dutch children were used. The first sample consisted of children aged 3 to 46 months. Details of this population are described in [107]. For the purpose of the current study, data on children aged 3.0 to 3.11 years of age were used (n=94; 45% boys). The second sample consisted of school children aged 4 to 13 years (see

---

1   All children being preschoolers, with a maximum age analogue to the maximum age for the Dutch adaptation of the Reynell Developmental Language Scales.

[108] for a detailed description of this population). From this sample, data on children aged 5.0 to 6.11 years were selected (n=378; 52% boys).

Health-related quality of life measures

During baseline measurement, a questionnaire booklet was filled in by the caregiver. Among other questionnaires, these booklets contained proxy versions of three different types of validated HRQOL measures. The Child Health Questionnaire was applied to provide a generic descriptive profile of HRQOL on distinct domains relevant for children. The Health Utility Index 3 also provides a description of health status and HRQOL but was specifically used because it is a utility (preference-) based measure, thereby able to summarize overall HRQOL in a single score. The EuroQoL EQ-VAS was also incorporated to indicate overall health in a single value. Presented below is a brief summary of the applied instruments.

Child Health Questionnaire

The Child Health Questionnaire (CHQ) is a family of generic, multidimensional HRQOL descriptive instruments for measurement in children. In our study, the Infant and Toddler Quality of Life Questionnaire (ITQOL-97) was used for children aged 3 and 4 years, and the Child Health Questionnaire - Parent Form 28 (CHQ-PF28) for children aged 5 and 6 years.

Infant and Toddler Quality of Life Questionnaire (ITQOL-97)

The ITQOL-97 [107,109] is a HRQOL instrument which has been translated into Dutch following international standards and validated in a general population sample in the Netherlands [107]. The ITQOL has good internal consistency, with all Cronbach's alpha > 0.70. Test-retest ICCs were moderate or adequate (≥ 0.50; $p<0.01$) and concurrent and discriminative validity has been shown [107]. The questionnaire consists of 97 items covering the following eight concepts of HRQOL in children aged 2 months to 5 years: *physical functioning, growth and development, bodily pain and discomfort, temperament and moods, general behavior, getting along with others, general health perceptions and change in health.* In addition, three parent-focused concepts are included: anxiety and worry due to the child's health (*parental impact: emotional*), limitations in time to meet parents own needs as a result of their child's health (*parental impact: time*) and how well one's family gets along with each other (*family cohesion*). Following the ITQOL scoring manual, the item scores are summed up and transformed to scale scores ranging from 0 (worst possible health state) to 100 (best possible health state). Thus, lower scores correspond to lower HRQOL.

2

Child Health Questionnaire – Parent Form 28 (CHQ-PF28)

The CHQ-PF28 [108,110] is a shortened version of the CHQ-PF50. Score distributions and discriminative validity are comparable to its longer counterpart [108]. The questionnaire consists of 28 items measuring nine HRQOL concepts in children aged 5 to 18: *physical functioning*, *role functioning: emotional/behavior*, *role functioning: physical*, *bodily pain*, *general behavior*, *mental health*, *self-esteem*, *general health perceptions* and *change in health*. The same three parent-focused concepts as in the ITQOL-97 are also included (*parental impact: emotional*, *parental impact: time*, *family cohesion*), and besides these, the concept of limitations in time to do *family activities*. As in the ITQOL-97, raw item scores are summed up and scale scores are transformed to a 0-100 scale (0 representing worst health and 100 best health). Scores can be analyzed separately and can be combined to obtain an overall physical and psychosocial summary score. These overall scores are derived from weighted composites of scale scores. Summary scores of 50 represent the mean in a US reference population children; 10 points above or below 50 reflect one standard deviation difference in either direction.

Health Utility Index 3 (HUI3)

The HUI3 [111] is a generic preference-based measure of health for children aged 5 years and above. It provides a comprehensive, reliable, responsive and valid measurement of health status and HRQOL in clinical studies [112]. This instrument has two components: 1) a standardized descriptive system for describing health or its impact on HRQOL; and 2) an algorithm for assigning values (utilities) to each health state described by the system. The HUI3 incorporates eight domains: *vision*, *hearing*, *speech*, *ambulation*, *dexterity*, *emotion*, *cognition* and *pain and discomfort (not specified)*, with five to six response levels for each domain. In all domains, level 1 represents perfect health. Responses on other levels represent decrements in health states. The values attached to the various health states reflect preferences for health that are derived from the general population of adults, and range from -.36 (worst imaginable health) to 1 (best imaginable health) [112]. Because the HUI3 is applicable for children aged 5 years and above, in the current study only parents of children aged 5 and 6 years old filled in the HUI3.

EuroQoL EQ-VAS (EQ-VAS)

The EQ-VAS [113] is a visual analog scale for directly recording an individual's rating for his or her current health state. In this study a proxy version was applied: Parents were asked to rate their child's current health state on a scale ranging from 0 (worst imaginable health) to 100 (best imaginable health). A proxy version of the EQ-VAS has been previously applied in pediatric populations, for instance in children with attention-deficit hyperactivity disorder [114] and children with chronic arthritis [115] and showed good validity properties in these populations. The EQ-VAS was used for all participating children.

Clinical measures

Information on stuttering severity and TSO of stuttering was obtained from the initial assessment by the speech- language pathologists. For the current study, the Stuttering Severity Index-3 (SSI-3) [116] was used as an index of stuttering severity. The stuttering severity score based on the SSI-3 is a weighted mean of (a) stuttering frequency, (b) duration of the three longest stutters and (c) physical concomitants. The following stuttering severity categories were distinguished for the study: (1) mild stuttering (SSI-3 score: 11-16)[2]; (2) moderate stuttering (SSI-3 score: 17-26); severe or very severe stuttering (SSI-3 score: 27 and above). In the RESTART-trial, the speech-language pathologists calculated the SSI-3 severity category based on video recordings in the clinic. Since it is difficult for parents (and often impossible) to recall the exact date of onset of stuttering, TSO was divided in three categories (in line with [34]): 6-12 months, 13-18 months and 19 months or longer.

Analysis

First, the distribution of scores for all HRQOL instruments were separately compared with the distribution of HRQOL values of the reference groups. Since most of the data distributions were severely skewed to the left, comparisons were based on the non-parametric Mann-Whitney U Test. The global level of significance was defined at $α = 0.05$. To correct for multiple comparisons per questionnaire, the Bonferroni correction was performed. As 11 different scores were compared for the ITQOL-97, the level of significance was defined at $α = 0.05/11 = 0.0045$. For the instruments applied in children aged 5-6 years, $α$ was set at $0.05/15 = 0.0033$ (CHQ-PF28) and $0.05/9 = 0.0055$ (HUI3). In order to indicate the clinical significance of observed statistically significant differences, effect sizes (ES) were calculated by dividing the difference in mean scores between CWS and reference groups by the pooled standard deviation. A small effect is indicated by an ES of 0.2 until 0.5, a moderate effect by an ES of 0.5 until 0.8 and a large effect by an ES of 0.8 and higher [117]. Second, analysis of covariance (ANCOVA), adjusting for age in months, was conducted for group comparisons in the study population of CWS. The between-subject factors were SSI (3 levels) and TSO (3 levels). Eta squared values were obtained to measure the effect sizes. Post-hoc tests (Bonferroni) were used to perform pairwise comparisons among groups. This was done for all ITQOL-97 and CHQ-PF28 scales and for the HUI3 utility and EQ-VAS score as dependent variables. Preliminary checks were conducted to ensure that assumptions of ANCOVA were not violated. For a few CHQ-PF28 outcomes, the assumption of homogeneity of variances was not met. Therefore, for these scales analyses were also conducted with logarithmic transformations of HRQOL scores, which did not change any of the results. All data were analyzed using SPSS 19.0 (Armonk, NY: IBM Corp.).

---

2   Children with a stuttering frequency < 3% in the therapy setting and ≥ 3% in the home setting were included in the group "mild stuttering"

RESULTS

Table 2.1 shows the clinical characteristics of the study population. Table 2.2 presents the mean scale scores on the ITQOL-97 and the CHQ-PF28 of the study population and the reference groups, and the *p*-value and effect-size for each scale. CWS scored higher on the *getting along* scale of the ITQOL-97, but after Bonferroni correction this difference was no longer significant. On all other scales no significant differences in scores between CWS and reference groups were found. Table 2.3 shows the mean HUI3 and EQ-VAS values for CWS and reference groups. CWS aged 5-6 years had lower mean scores for the HUI3 domains *speech*, *emotion* and *cognition;* however, after Bonferroni correction only the *speech* domain remained significant. The effect size was large (0.99 vs. 0.95, original *p* < .01, Cohen's *d* = 1.0). The HUI3 utility reflected a lower HRQOL for 5-6 year old CWS, revealing a moderate effect size (0.94 vs. 0.88, original *p* < .01, Cohen's *d* = 0.60). The EQ-VAS for all CWS showed a small to moderate effect size (92.40 vs. 88.14, original *p* < .01, Cohen's *d* = 0.45). A separate analysis of EQ-VAS scores for CWS aged 3-4 years and 5-6 years compared with reference scores revealed similar results.

TABLE 2.1. Stuttering characteristics of participating CWS

Characteristics	Total group (N = 197)	Participants 3-4 years (N = 146)	Participants 5-6 years (N = 51)
Stuttering severity			
Mild	64 (32.5%)	48 (32.9%)	16 (31.4%)
Moderate	97 (49.2%)	72 (49.3%)	25 (49.0%)
Severe	36 (18.3%)	26 (17.8%)	10 (19.6%)
Time since onset			
6-12 months	87 (44.2%)	79 (54.1%)	8 (15.7)
13-18 months	48 (24.4%)	35 (24.0%)	13 (25.5)
19 months or longer	62 (31.5%)	32 (21.9%)	30 (58.8)

Figure 2.1 shows the distribution of responses on the HUI3 domains for the study population of stuttering children aged 5-6 years and for the reference population. Compared with the reference population, parents of CWS more often chose level 2 and 3 of the speech domain (representing lower scores on this domain of health). For the domains *emotion* and *cognition*, the difference was mainly expressed in more level 2 problems in stuttering children.

TABLE 2.2 Results of the ITQOL-97 and CHQ-PF28 for CWS and reference groups

	Reference group[b]			CWS 3-4 years			p-value	Effect size (Cohen's d)
	N	Mean	SD	N	Mean	SD		
ITQOL-97 scale								
Physical functioning	91	98.86	2.80	144	97.98	10.09	.21	0.14
Growth and development	94	85.78	11.01	145	83.23	12.90	.11	0.21
Bodily pain and discomfort	93	88.17	15.94	145	90.43	11.99	.45	0.16
Temperament and moods[a]	94	76.78	9.55	67	76.45	9.05	.94	0.04
General behavior[a]	94	72.21	12.62	65	73.96	13.47	.56	0.13
Getting along	94	71.95	8.44	145	74.72	8.86	.01*	0.32
General health perceptions	94	78.52	15.18	145	82.18	13.11	.07	0.26
Change in health	94	56.65	18.76	144	58.51	20.09	.60	0.10
Parental impact: emotional	94	90.12	10.70	145	90.76	10.08	.79	0.06
Parental impact: time	94	93.47	12.91	145	93.71	10.47	.72	0.02
Family cohesion	94	73.24	18.99	142	75.81	16.23	.42	0.15

	Reference group[c]			CWS 5-6 years			p-value	Effect size (Cohen's d)
	N	Mean	SD	N	Mean	SD		
CHQ-PF28 scale								
Physical functioning	378	96.30	12.38	46	96.86	13.99	.51	0.04
Role functioning-emotional/behavior	378	96.74	12.31	47	95.74	13.22	.49	0.08
Role functioning-physical	378	95.59	14.30	47	95.74	13.22	.92	0.01
Bodily pain	378	81.22	16.53	49	84.49	14.87	.19	0.21
General behavior	378	69.78	15.23	48	71.94	13.65	.38	0.15
Mental health	378	81.77	13.66	49	79.76	16.14	.53	0.13
Self-esteem	378	79.92	14.54	48	78.82	12.03	.51	0.08
General health perceptions	378	85.22	16.47	49	84.28	17.70	.88	0.06
Change in health	378	55.36	16.27	49	56.12	14.90	.56	0.05
Parental impact: emotional	378	85.52	15.72	49	85.20	15.24	.83	0.02
Parental impact: time	378	93.47	13.08	47	92.20	18.67	.99	0.08
Family cohesion	378	69.88	18.93	49	74.49	19.59	.10	0.24
Family activities	378	86.74	17.58	48	88.54	17.07	.41	0.10
Physical summary score	378	55.74	7.45	44	56.01	7.53	.74	0.04
Psychosocial summary score	378	52.01	7.13	44	51.72	5.86	.40	0.04

[a] Due to a failure in the data collection, analyses for these subscales were based on a smaller sample size of CWS
[b] Reference population from Raat et al. [107]: children 3.0-3.11 years of age
[c] Reference population from Raat et al. [108]: children 5.0-6.11 years of age
* p < .05, significant before Bonferroni correction
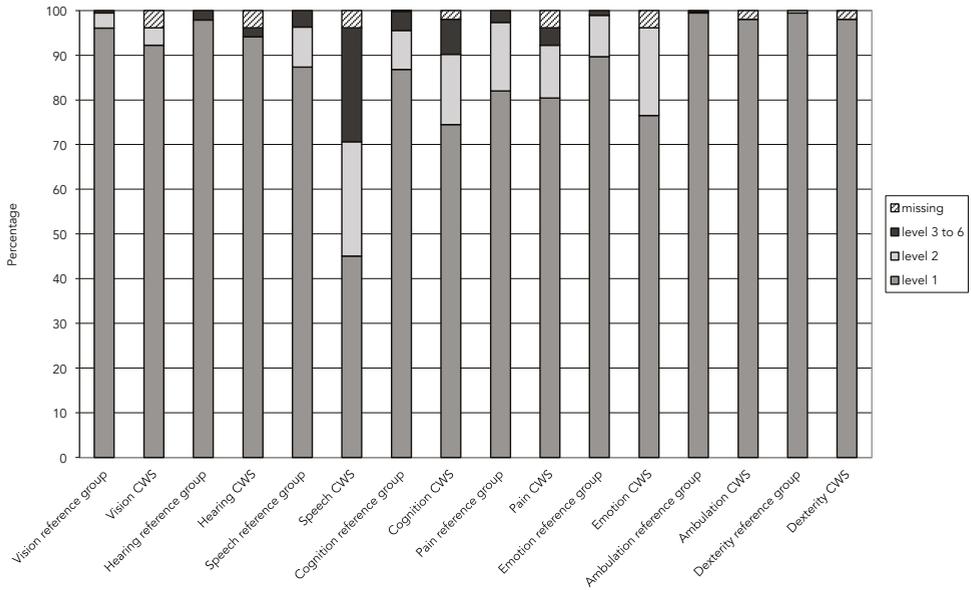
FIGURE 2.1 Percentage distribution of responses by HUI3 domain for CWS and reference group

TABLE 2.3 Results of the HUI3 and EQ-VAS for CWS and reference groups

	Reference group[a]			CWS 5-6 years			p-value	Effect size (Cohen's d)
	N	Mean	SD	N	Mean	SD		
HUI3 dimension								
Vision	378	1.00	0.01	49	1.00	0.00	.98	0.00
Hearing	378	1.00	0.01	49	1.00	0.03	.90	0.00
Speech	378	0.99	0.03	49	0.95	0.05	<.01**	1.00
Ambulation	378	1.00	0.01	49	1.00	0.00	.61	0.00
Dexterity	378	1.00	0.00	49	1.00	0.00	.61	0.00
Emotion	378	0.99	0.02	49	0.99	0.02	.04*	0.00
Cognition	378	0.99	0.04	49	0.98	0.04	.05*	0.25
Pain	378	0.99	0.02	49	0.99	0.02	.81	0.00
HUI3 multi-attribute utility score	378	0.94	0.09	49	0.88	0.11	<.01**	0.60
	Reference group[b]			All CWS			p-value	Effect size (Cohen's d)
	N	Mean	SD	N	Mean	SD		
EQ-VAS	450	92.40	8.85	183	88.14	10.14	<.01**	0.45

[a] Reference population from Raat et al. (2005): children 5.0-6.11 years of age
[b] Reference population from Raat et al. (2005): children 4.0-6.11 years of age
* p < .05, significant before Bonferroni correction
** p < .0055, significant after Bonferroni correction (α=.05/9=.0055)

ANCOVA-analyses in children aged 3-4 years showed statistically significant main effects for stuttering severity for the ITQOL-97 scales *temperament and moods* and *family cohesion* (see Table 2.4). However, effect sizes were very small (eta squared=0.001 and 0.002 respectively). Post-hoc tests showed that on both scales children with mild stuttering scored higher (indicating better health) than children with moderate and/or severe stuttering (see Table 2.5), but differences were not statistically significant.

TABLE 2.4 ANCOVA results

Dependent variable	df	F	*p*-value	Eta squared
Independent variable = stuttering severity				
ITQOL-97 temperament and moods	2, 61	3.56	.035	0.001
ITQOL-97 family cohesion	2, 136	3.11	.048	0.002
CHQ-PF28 general health perceptions	2, 43	4.34	.019	0.007
CHQ-PF28 physical summary score	2, 38	3.63	.036	0.003
Independent variable = TSO				
CHQ-PF28 physical functioning	2, 40	3.40	.043	0.003
CHQ-PF28 family cohesion	2, 43	4.95	.012	0.011

Note. Only significant effects are displayed

For children aged 5-6 years, statistically significant main effects for stuttering severity were found for the CHQ-PF28 scales *general health perceptions* and *physical summary score*, with very small effect sizes (eta squared=0.007 and 0.003, respectively; see Table 2.4). On both scales, children with mild stuttering had significantly higher scores (indicating better health) than children with moderate stuttering (see Table 2.5). On the scales *physical functioning* and *family cohesion*, statistically significant main effects were found for TSO. Effect sizes were again very small (eta squared= 0.003 and 0.011, respectively; see Table 2.4). Post-hoc tests revealed higher scores for children who stuttered 19 months or longer, compared with children who stuttered 13-18 months (see Table 2.5).

On the EQ-VAS and HUI3 no effect of stuttering severity or TSO was established. HRQOL results for groups of CWS with different degrees of stuttering severity and different TSO are presented in respectively Appendix A2.1 and A2.2.

TABLE 2.5 Adjusted mean ± standard error of the ITQOL-97 and CHQ-PF28 subscale scores by group of stuttering severity and TSO

	Stuttering severity			Time since onset		
	Mild stuttering	Moderate stuttering	Severe stuttering	TSO 6-12 months	TSO 13-18 months	TSO 19+ months
ITQOL-97 temperament and moods	82.42 (2.29)	76.61 (1.59)	74.95 (2.70)			
ITQOL-97 family cohesion	81.37 (2.55)	74.53 (1.97)	73.18 (3.23)			
CHQ-PF28 general health perceptions	89.60* (4.52)	74.34 (3.81)	85.89 (5.36)			
CHQ-PF28 physical summary score	59.61* (2.11)	52.84 (1.71)	55.33 (2.30)			
CHQ-PF28 physical functioning				99.64 (4.84)	87.65 (4.32)	100.38* (2.68)
CHQ-PF28 family cohesion				65.24 (6.99)	66.03 (5.36)	82.65* (3.53)

Note. Only results for significant group differences in the ANCOVA-analysis are displayed. Outcomes for all HRQOL scales for CWS in different groups are presented in the Appendix

* $p <.05$

DISCUSSION

This study explored the HRQOL of 197 preschool stuttering children referred for treatment using three validated generic instruments. The outcomes on the EQ-VAS for all preschool children and the HUI3 utility value (only available for children aged 5-6 years old) showed that the HRQOL of CWS is slightly poorer than that of the norm in the Dutch population (small to moderate effect sizes), but results on the different health domains of the CHQ-instruments (available for children aged 3-6 years old) did not reveal any difference in scores between stuttering children and reference groups. Within the group of CWS, two individual ITQOL-97 and four CHQ-PF28 scales showed statistically different scores for children with different SSI stuttering severity or TSO. However, based on the effect sizes, these differences were considered negligible. Thus, our study did not establish a meaningful effect of severity or TSO on the HRQOL of preschool CWS.

Interestingly, the results of our study confirm the conclusion of the recent population-based study by Reilly et al. [11]. This study suggested no general negative impact of developmental stuttering on the HRQOL of young children, even though the range of responses in the study indicated that the HRQOL of a few individual stuttering children was impaired. If a general HRQOL impairment were present in preschool CWS, this would have been likely to emerge in the data from our clinical population of stuttering children. All children in our study had stuttered for at least six months and their stuttering had

been confirmed by clinical measures, so no ambiguous or borderline cases were included. Furthermore, parents who seek help for a health problem for their child can be expected to be *concerned* about their child; it is typically parents' perception of their children's HRQOL that influences health care utilization [118]. However, the CHQ-instruments used in our study did not show more parental concerns due to the child's health in parents of CWS compared to parents in the reference populations.

Exceptions to the overall null-finding appear to be the results of our population of CWS showing a significantly lower mean HUI3 speech domain score (with a large effect size) and EQ-VAS score (small to moderate effect size), compared to the reference populations. This finding of a lower score on the HUI3 speech domain (representing diminished functioning in this domain), is not surprising in a population of CWS, since it could be interpreted as an acknowledgment of parents judging the speech of their children to be less favorable than parents of non-stuttering children. The finding of the lower EQ-VAS score cannot be easily interpreted; however, it is important to note that the EQ-VAS does not measure HRQOL through the degradation of health on distinct domains, but rather asks directly for a (proxy) HRQOL value.

Since, in general, demands on the oral communication skills of young children are limited, impaired speech may not critically influence their daily functioning. Furthermore, very young children, particularly if their stuttering is mild, may hardly be aware of their stuttering and negative reactions from peers may still be exceptional. Finally, emotional well-being in young children may depend less on the social environment and more on the trusting relationships with their parents. On the other hand, the lower score on the *temperament and moods* scale of the ITQOL-97 for CWS with moderate and severe stuttering compared to children with mild stuttering might suggest that young children who stutter severely have more difficulties in regulating their moods and temperaments (e.g., fussiness, sleeping difficulties, lack of alertness). This finding possibly indicates a connection between temperament and stuttering severity, which would be in line with results of several studies into these two entities in the last decade (for an overview, see [100]). However, the effect size in our study was too small to be interpreted as relevant. The same holds for the lower scores on the *family cohesion* scale scores for children with mild stuttering in this age group. With regard to the higher family cohesion score (again with a small effect size) for children aged 5-6 years who stuttered for more than 19 months compared to children who stuttered 13-18 months, one could speculate that families of these children had become closer, for instance to compensate for problematic relationships with peers or to protect or additionally support their stuttering child. Higher family cohesion for families of non-healthy children has also been reported by Bannink, Maliepaard, Raat, Joosten, and Mathijssen [119].

2

Physical problems were not a priori anticipated in our study group of CWS. Although our data suggest a relationship between physical dysfunctioning and stuttering, the effect sizes were too small to support a firm conclusion. Besides, this association may also have been caused by the coincidental presence of more children with physical problems in the groups of CWS aged 5-6 years with moderate stuttering and the group who stuttered for 13-18 months. A review of the available data on the physical state of CWS aged 5-6 years old confirms that relatively many children with moderate or severe stuttering (n=7; 14%) were being checked by a doctor regularly and/or were using medicines (e.g., for eczema or otitis media), compared with none of the children with mild stuttering.

Since this was the first study of its kind, there are a number of limitations which need to be considered. Firstly, different HRQOL instruments had to be used for different age groups. This hampered a direct comparison of results across all age groups. The EQ-VAS was the only instrument that could be used for all children. Splitting the groups also resulted in relatively small groups; the group with severe stuttering in children aged 5-6 years was particularly small (n=10). Therefore, in some cases the conditions for establishing a statistical significant result (i.e., sufficient statistical power) could not be fully met, for instance with regard to the low mean score on the *mental health* domain of the CHQ-PF28. This also holds for the other domains. Since the current study was part of a larger study in which HRQOL was a secondary outcome measure, an a priori power calculation was not possible for this study. Nevertheless, the mostly small effect sizes found in our study do strengthen our conclusions on the null-finding.

Secondly, although a combination of generic and disorder-specific instruments is generally recommended in HRQOL research [88], currently no stuttering-specific HRQOL measure for young children exists. It might be argued that the incorporated HRQOL measures in the current measurements are not specific enough to detect a HRQOL impairment, even though they cover emotional and social domains of health. While the scale *general health perceptions* of the CHQ-instruments is known to be highly sensitive in somatic pediatric conditions, the sensitivity of the CHQ-instruments for conditions related to mental health has yet to be further investigated [120].

Thirdly, the reference data for the ITQOL-97 were only available for children aged 3.0-3.11 years, while our study population consisted of children aged 3.0-4.11 years. However, the fact that in our study population the scores did not differ between children aged 3 and 4 years, supports the validity of our conclusions.

Finally, the HRQOL results are based on parent-reported data. The use of parental proxies was essential considering the age of the children in this study. However, it is known that the adequacy of proxy ratings may be confounded by certain characteristics such as parental emotions and stress [121-123]. It was not possible to take these into consideration in the current study. Furthermore, correlations between child and proxy reports are known to be high for observable, physical domains but lower for less observable domains like pain and anxiety [124]. Thus, the impact of stuttering on the emotional and social HRQOL domains, which have been shown to be most affected in adults who stutter, might not be well estimated by proxy reports. However, a difference would not necessarily have to mean that the child reports a lower HRQOL. For instance a study into chronic pain in children showed the opposite, i.e., parents rated their child's HRQOL lower than the child himself [125]. The relationship between HRQOL ratings by stuttering children and proxies needs to be explored in future studies. Additional research incorporating child self-report during follow-up at school age would expand on our findings.

CONCLUSION

In conclusion, while stuttering that persists into adolescence and adulthood is commonly associated with a reduced HRQOL [41,42,84,85,91,92,126], the results of the current study do not reveal a diminished HRQOL for preschool stuttering children from a clinical population. Intervention during the preschool years may be crucial to try to prevent stuttering becoming a chronic, severe condition and thus potentially decrease an individual's well-being. To provide more insight into the broader health consequences of stuttering in children, we would recommend further research, which would include a larger cohort of children with severe stuttering and study the longitudinal course of HRQOL. Incorporating additional parameters, such as child functioning and characteristics of both the child (e.g., temperament and presence of social anxiety) and the environment (e.g., family climate and parental stress), could expand on this study and provide a more comprehensive interpretation of HRQOL results in children who stutter.

ACKNOWLEDGMENTS

APPENDIX

TABLE A2.1 Results on the ITQOL-97, CHQ-PF28, HUI3 and EQ-VAS in CWS for different stuttering severity groups

	Mild stuttering			Moderate stuttering			Severe stuttering		
	N	Mean	SD	N	Mean	SD	N	Mean	SD
ITQOL-97 item scale									
Physical functioning	48	99.79	1.08	71	96.28	14.15	26	99.36	1.64
Growth and development	48	84.95	11.20	71	82.72	10.71	26	81.44	19.78
Bodily pain and discomfort	48	92.71	10.12	71	89.38	12.67	26	89.10	13.07
Temperament and moods[a]	20	80.31	9.78	35	75.30	8.77	12	73.38	6.85
General behavior[a]	16	80.39	15.69	36	70.65	12.09	13	73.94	11.02
Getting along	48	75.30	10.07	71	73.35	7.82	26	77.41	8.76
General health perceptions	48	84.85	12.12	71	80.65	13.39	26	81.45	13.84
Change in health	48	58.33	19.52	70	58.21	20.73	26	59.62	20.10
Parental impact: emotional	48	92.71	9.26	71	89.54	11.09	26	90.52	8.26
Parental impact: time	48	96.00	7.27	71	92.28	11.57	26	93.41	11.90
Family cohesion	48	80.10	17.37	68	74.12	15.31	26	72.31	15.31
CHQ-PF28 item scale									
Physical functioning	15	100	0.00	21	93.65	20.36	10	98.89	3.51
Role functioning-emotional/behavior	15	100	0.00	22	92.42	17.61	10	96.67	10.54
Role functioning-physical	15	100	0.00	22	93.94	16.70	10	93.33	14.05
Bodily pain	16	90.00	10.33	23	83.48	16.68	10	78.00	14.76
General behavior	16	70.86	14.15	22	74.17	13.88	10	68.79	12.83
Mental health	16	80.21	17.18	23	83.70	12.68	10	70.00	18.92
Self-esteem	16	82.29	10.49	23	77.65	12.44	10	75.83	13.29
General health perceptions	16	91.88	13.09	23	77.54	19.55	10	87.63	14.83
Change in health	16	54.69	13.60	23	57.61	13.97	10	55.00	19.72
Parental impact: emotional	16	91.41	9.92	23	84.24	18.16	10	77.50	11.49
Parental impact: time	14	97.62	6.05	23	88.41	24.84	10	93.33	11.65
Family cohesion	16	73.13	22.50	23	72.83	19.35	10	80.50	15.36
Family activities	15	89.17	18.82	23	86.41	18.43	10	92.50	10.54
Physical summary score	13	59.58	1.69	21	53.50	9.87	10	56.62	4.14
Psychosocial summary score	13	53.70	4.72	21	52.12	5.54	10	48.30	6.88
HUI3 dimension									
Vision	15	1.00	0.01	24	1.00	0.00	10	1.00	0.00
Hearing	15	1.00	0.00	24	0.99	0.04	10	1.00	0.00
Speech	15	0.96	0.04	24	0.96	0.04	10	0.94	0.05
Ambulation	16	1.00	0.00	24	1.00	0.00	10	1.00	0.00
Dexterity	16	1.00	0.00	24	1.00	0.00	10	1.00	0.00
Emotion	15	0.99	0.02	24	0.99	0.02	10	0.98	0.03
Cognition	16	0.98	0.05	24	0.98	0.04	10	0.98	0.03
Pain	15	1.00	0.00	24	0.99	0.03	10	0.99	0.03
HUI3 multiattribute utility score	15	0.90	0.12	24	0.88	0.11	10	0.86	0.11
EQ-VAS	62	88.94	9.89	88	87.42	10.93	33	88.58	8.44

[a] Due to a failure in the data collection, analyses for these subscales were based on a smaller sample size of CWS

Table A2.2 Results on the ITQOL-97, CHQ-PF28, HUI3 and EQ-VAS in CWS for different time since onset (TSO) groups

	TSO 6-12 months			TSO 13-18 months			TSO ≥ 19 months		
	N	Mean	SD	N	Mean	SD	N	Mean	SD
ITQOL-97 item scale									
Physical functioning	77	97.45	12.69	35	98.74	6.89	32	98.44	4.56
Growth and development	78	82.46	14.43	35	85.70	11.49	32	82.40	10.04
Bodily pain and discomfort	78	91.45	12.23	35	91.07	11.08	32	87.24	12.16
Temperament and moods [a]	44	75.63	9.55	9	81.64	7.54	14	75.69	7.58
General behavior [a]	38	72.11	12.65	12	75.71	17.83	15	77.23	11.57
Getting along	78	74.22	8.82	35	76.02	9.30	32	74.51	8.59
General health perceptions	78	82.33	12.57	35	83.18	13.22	32	80.72	14.54
Change in health	78	60.26	20.72	34	56.62	17.74	32	56.25	21.06
Parental impact: emotional	78	90.61	10.92	35	90.61	9.92	32	91.29	8.21
Parental impact: time	78	93.76	10.84	35	93.02	10.04	32	94.35	10.29
Famlily cohesion	78	74.94	16.58	34	76.91	17.84	30	76.83	13.55
CHQ-PF28 item scale									
Physical functioning	8	98.61	3.93	10	87.78	28.90	28	99.60	2.10
Role functioning-emotional/behavior	8	95.83	11.79	10	90.00	22.50	29	97.70	8.60
Role functioning-physical	8	100.00	0.00	10	90.00	22.50	29	96.55	10.33
Bodily pain	8	82.50	16.69	12	88.33	13.37	29	83.45	15.18
General behavior	8	68.91	12.47	11	66.59	15.30	29	74.81	12.96
Mental health	8	79.17	14.77	12	73.61	20.67	29	82.47	14.15
Self esteem	8	83.33	12.60	11	79.55	8.63	29	77.30	12.97
General health perceptions	8	80.00	22.64	12	78.30	20.16	29	87.93	14.67
Change in health	8	56.25	11.57	12	58.33	22.19	29	55.17	12.28
Parental impact: emotional	8	79.69	11.45	12	83.33	16.28	29	87.50	15.67
Parental impact: time	8	95.83	7.72	11	83.42	30.46	28	94.64	13.65
Family cohesion	7	68.57	23.22	12	63.33	19.46	30	80.33	16.86
Family activities	8	87.50	11.57	12	87.50	19.22	28	89.29	17.91
Physical summary score	8	56.43	4.34	9	52.56	14.13	27	57.03	4.69
Psychosocial summary score	8	51.25	6.46	9	50.18	6.35	27	52.37	5.65
HUI3 dimension									
Vision	8	1.00	0.01	11	1.00	0.00	30	1.00	0.00
Hearing	8	1.00	0.00	11	1.00	0.00	30	0.99	0.04
Speech	8	.94	0.05	11	0.97	0.05	30	0.95	0.05
Ambulation	8	1.00	0.00	12	1.00	0.00	30	1.00	0.00
Dexterity	8	1.00	0.00	12	1.00	0.00	30	1.00	0.00
Emotion	8	0.98	0.03	11	0.99	0.02	30	0.99	0.02
Cognition	8	0.95	0.04	12	0.97	0.07	30	0.99	0.02
Pain	8	0.99	0.02	11	0.98	0.03	30	0.99	0.02
HUI3 multiattribute utility score	8	0.82	0.14	11	0.87	0.14	30	0.90	0.09
EQ-VAS	82	87.27	10.70	42	88.56	9.22	58	89.07	9.22

[a] Due to a failure in the data collection, analyses for these subscales were based on a smaller sample size of CWS

3

Health-related quality of life of adults who stutter

With Clazien Bouwmans
Marie-Christine Franken
Elly Stolk

ABSTRACT

Although persistent developmental stuttering is known to affect daily living, just how great the impact is remains unclear. Furthermore, little is known about the underlying mechanisms which lead to a diminished quality of life (QOL). The primary objective of this study is to explore to what extent QOL is impaired in adults who stutter (AWS). In addition, this study aims to identify determinants of QOL in AWS by testing relationships between stuttering severity, coping, functioning and QOL and by testing for differences in variable scores between two AWS subgroups: receiving therapy versus not receiving therapy. A total of 91 AWS filled in several questionnaires to assess their stuttering severity, daily functioning, coping style and QOL. The QOL instruments used were the Health Utility Index 3 (HUI3) and the EuroQoL EQ-5D and EQ-VAS. The results indicated that moderate to severe stuttering has a negative impact on overall quality of life; HUI3 derived QOL values varied from .91 (for mild stuttering) to .73 (for severe stuttering). The domains of functioning that were predominantly affected were the individual's *speech*, *emotion*, *cognition* and *pain* as measured by the HUI3 and *daily activities* and *anxiety/depression* as measured by the EQ-5D. AWS in the therapy group rated their stuttering as more severe and recorded more problems on the HUI3 speech domain than AWS in the non-therapy group. The EQ-VAS was the only instrument that showed a significant difference in overall QOL between groups. Finally, it was found that the relationship between stuttering severity and QOL was influenced by the individual's coping style (emotion-oriented and task-oriented). These findings highlight the need for further research into stuttering in relation to QOL, and for a broader perspective on the diagnosis and treatment of stuttering, which would take into consideration quality of life and its determinants.

3

INTRODUCTION

How stuttering affects the overall quality of life (QOL) of adults who stutter (AWS) has not yet been extensively researched. This is surprising since about 1% of the adult population stutters [7] and because it is known that AWS often experience negative affective, behavioral, and cognitive reactions. Moreover, stuttering significantly limits the speaker's ability to participate in daily activities [86]. AWS frequently experience disabling levels of social anxiety [127-129]. Whether this occurs, depends on their fear of a negative evaluation in social relations because of their stuttering, and whether or not they act upon that fear by adoption of a strategy of avoidance [130]. Recognizing the complexity of the stuttering disorder in adults, researchers have established the need to document not only speech symptoms, but also broad-based outcome parameters such as QOL (e.g., [41,75,78,86]). Recently, a special edition of the Journal of Fluency Disorders dedicated to the QOL of people who stutter also raised awareness for the topic [9,77,131]. Until now, the magnitude of and mechanisms underlying the QOL effects of stuttering in adults have not been fully explored. The purpose of this study is to evaluate QOL in AWS by means of a comprehensive assessment.

A review of existing literature with respect to the QOL in AWS revealed that the majority of studies use a narrow conceptualization of QOL. That is, most studies investigated the QOL of AWS by focusing on the influence that stuttering has on the specific life domains which are believed to be most affected by stuttering (e.g., [42,83-85,132]). For example, Hayhow et al. [84] showed the major adverse effects of stuttering on school life and occupational choice. The negative impact of stuttering on school performance, relationships with teachers and classmates, and performance at work was confirmed by Klompas and Ross [85], who interviewed 16 AWS. Klein and Hood [42] found that the majority of the AWS perceived their stuttering to be a handicap in relation to employment opportunities and job performance. By exploring specific life domains potentially affected by the condition stuttering, these studies provide significant, but only limited, information on QOL. A disadvantage of such *condition-specific* QOL studies is that little insight is gained into the *overall* QOL (e.g., [41,75,133]). In other words, although these studies provide insight into problems associated with stuttering, not all aspects of QOL relevant to a person are taken into account. In addition, due to the incorporation of dissimilar domains, condition-specific QOL instruments cannot be used to compare different health conditions.

In contrast to condition-specific QOL instruments, generic QOL instruments embrace a broad conceptualization of QOL by measuring a comprehensive set of domains. A common element in these generic QOL instruments is the incorporation of physical, emotional and social domains of health. These domains are relevant for anyone, irrespective of the

specific health problem. As a result, generic QOL instruments are suitable for comparison of stuttering to other health states. Well-known examples are the Medical Outcomes Study Short Form 36-Item Health Survey (SF-36) and the Nottingham Health Profile (NHP). A limitation of these descriptive generic QOL instruments is that they do not quantify how each dimension contributes to overall well-being. That is, if some domains are significantly affected but others are not, the effect on overall QOL cannot be established. To overcome this problem, QOL researchers frequently move beyond a multidimensional generic description of health by attaching a single value to the overall health status [133]. This value or *utility* summarizes all the positive and negative aspects of health into one single QOL index, which is usually set between 0, which corresponds to a health state valued as equivalent to death, and 1, which corresponds to perfect health. Such QOL values can be established in two ways. Firstly, health states can be estimated by using validated *preference-based* techniques [134]: Visual Analogue Scale (VAS), Time Trade-Off (TTO) or Standard Gamble (SG). Alternatively, a special class of generic QOL instruments can be used, for which QOL values are available for all health states described by the instrument [135]. Well known examples are the EQ-5D, Health Utilities Index (HUI) and the SF-6D (derived from the SF-36) [136].

So far, only two studies have attempted to gain insight into overall QOL of AWS by using generic QOL instruments or by preference-based techniques. Craig et al. [41] used the SF-36 to explore the negative impact caused by stuttering in a population of AWS and adults who do not stutter (AWNS). The authors showed that, compared to a non-stuttering control group, stuttering affects social and emotional functioning, as well as vitality and mental health status. The effect sizes (standardized mean difference between the groups) on these domains varied between .28 and .59, indicating small to moderate QOL impairments in AWS. Because the associated SF-6D utilities were not reported by Craig et al. [41], the effect on overall QOL remains unclear. The study by Bramlett, Bothe and Franic [137] is the only study that we are aware of that adopted a preference-based approach to estimate utilities. Bramlett et al. [137] obtained overall QOL values for mild, moderate and severe stuttering from 75 AWNS using the three validated preference-based techniques mentioned before: VAS, TTO and SG. The results suggested that QOL is negatively affected by stuttering. Using the TTO method, non-stuttering adults valued their own health at .98 (SD .07), while they rated mild, moderate and severe stuttering at respectively .93 (SD .14), .85 (SD .18) and .63 (SD .24) [137]). Considering that a QOL weight of .63 has been found for living with home dialysis (Sackett and Torrance, 1978 in [137]), the results suggest that severe stuttering has a substantial impact on a person's overall QOL. However, the differences in QOL values between methods were substantial: VAS and SG resulted in QOL values of .44 (SD .20) and .81 (SD .19) respectively for severe stuttering. In addition, the applied methods provided little or no insight into the determinants (i.e., the underlying mechanisms) that lead to QOL impairments.

Thus the purpose of this study is to explore to what extent overall QOL is impaired in AWS and to investigate the determinants of such QOL impairment. Based on the conceptualization of QOL by Wilson and Cleary [79], the determinants measured in this study are stuttering severity, functioning and coping. Differences in stuttering severity, functioning, coping and QOL are examined between AWS who were in therapy and those who were not. Both groups are included because we hypothesize that studying QOL solely in a clinical population might lead to observing a greater reduction in QOL than when also taking into account the QOL of AWS who do not seek therapy. In terms of impaired quality of life, AWS who present themselves to a clinic might be those who are most severely affected by their condition. This could either be because their level of stuttering is more severe or because they have poorer coping skills and are more bothered by the effect of stuttering on their social interactions. Busschbach, Rikken, Grobbee, De Charro, and Wit [138] observed a lower QOL for adults with a short stature who had presented themselves to a clinic compared with a population based sample of short adults. It is thus considered important to include both AWS who were in therapy and AWS who were not in order to account for variability in the determinant variables of QOL in both groups and to provide a broad view of QOL. These insights could provide valuable support in designing possible starting-points for diagnosis, therapy, and measuring end points in clinical trials [139,140].

CONCEPTUALIZATION OF QOL

The empirical evaluation of QOL in AWS in this study is based on the theoretical conceptualization of QOL published by Wilson and Cleary [79]. This conceptual framework shows how different health measures can be combined to constitute a *broad* assessment of QOL. This section will explain how the QOL model extracted from the original Wilson and Cleary [79] model is built up.

The core of the model (Figure 3.1) is the relationship between symptoms, functioning and general health perception, the latter often referred to as health-related quality of life (HRQOL) or briefly as QOL (in this paper). *Symptoms* are defined as perceptual judgments of an abnormal physical, emotional, or cognitive state. *Functioning* refers to the ability of the individual to perform particular defined tasks. Basic domains of functioning that are commonly measured are physical, social, role and psychological functioning [135]. By measuring functioning on generic domains, the impact of a condition can be assessed in terms that are relevant to any individual. *General health perception* or *(Hr)QOL* reflects an overall, subjective evaluation of health status, in relation to symptoms and functional problems.

The model highlights the direct and indirect relationships between the adjacent outcome levels (how symptoms impact on functioning, and functioning on QOL), which can be assessed using condition-specific and generic outcome measures. In addition, the model takes into account that *characteristics of the individual* as well as *characteristics of the environment* might impact on the experience of symptoms, daily functioning and QOL and their relationships. These factors may have a direct or indirect impact on QOL. Examples of individual characteristics that affect QOL are psychological characteristics, personality and individual expectations. Examples of environmental characteristics are social support and the employment environment. The model does not precisely prescribe which of these factors may be relevant for exploration of QOL.
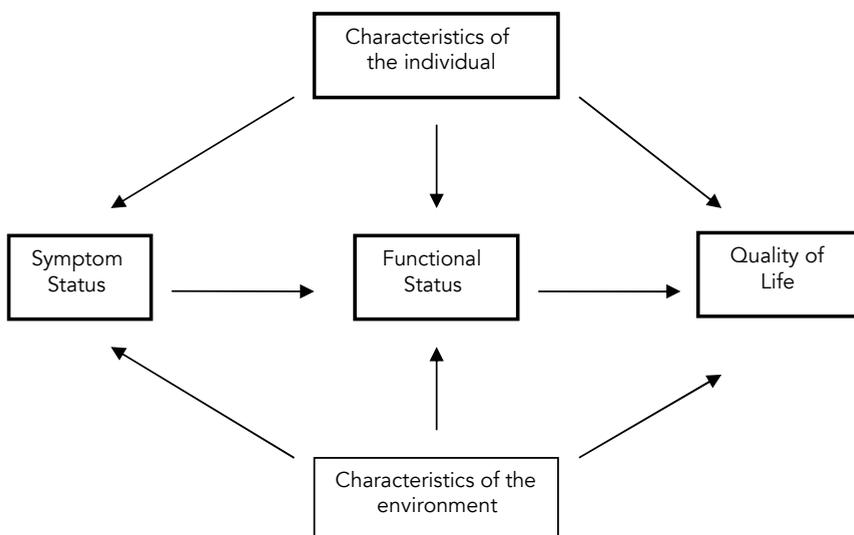
FIGURE 3.1 Conceptualization of determinants of quality of life, adapted from "Linking clinical variables with health-related qualify of life. A conceptual model of patient outcomes" [79]

OPERATIONALIZATION OF QOL IN AWS

This section indicates and motivates the selection of instruments for the operationalization of the Wilson and Cleary [79] model in this study. The instruments will be described in more detail in Section 4.3.

Overall QOL

To address the first aim of this study, to measure to what extent overall QOL is impaired in AWS, we assessed the primary outcome level (Quality of Life) by measuring QOL values for the health states reported by the participants. As described in Section 1, these values can

be obtained by using generic QOL instruments for which QOL values for all health states are available. We decided to use two widely applied the Health Utility Index 3 (HUI3; [111], the EuroQoL EQ-5D and the EuroQoL EQ-VAS [113]). These are brief and easy to use self-completed questionnaires.

Symptom status

The symptom status level can be assessed by measuring the level of stuttering severity experienced (e.g., [141]). In the current study, the level of stuttering severity was rated using two self-assessment scales. Self-assessment scales have shown to be correlated well with objective stuttering measures and other self-evaluation instruments [142] and rating by speech-language pathologists [141]. In addition to the self-assessment scales, a comprehensive stuttering instrument was applied; the Overall Assessment of the Speaker's Experience of Stuttering for adults (OASES, [86]).

Functioning status

In line with the Wilson and Cleary [79] model, functioning status was also measured in a generic way. The functional profiles provided by the HUI3 and EQ-5D were used as indicators of functioning. The domains that are measured by these instruments were considered as potentially relevant with regard to stuttering. That is, functioning restraints in AWS could be expected in the social, role and psychological domains, for instance communication in social situations or at work (e.g., [86]).

Characteristics of the individual

With respect to the individual and environmental characteristics, various studies suggest that coping style is an important determinant in QOL in AWS (e.g., [83,143-145]. Coping refers to the conscious response or reaction to events that are perceived as stressful [146]. An association between coping and QOL in AWS may be expected because individuals can adopt different strategies to reduce stress levels caused by their diminished ability to speak fluently. These coping styles may differ in their effectiveness to prevent negative QOL effects. Coping models have been frequently used to explain successful adjustment to chronic diseases, by showing the active role that patients may exert in managing the challenges that emanate from their condition [147]. Stuttering might well be a condition for which the applied coping style strongly influences the experienced QOL, since AWS are frequently confronted with their speech limitations [148]. In the current study, coping style was analyzed using the *Coping Inventory for Stressful Situations* (CISS) [149]. Environmental characteristics were not explicitly measured in this study, since it was argued that the main environmental factors related to QOL in AWS are also related to coping. That is, the social environment can be perceived as more or less demanding with regard to fluent speech and

therefore influence coping ability, and, adversely, coping styles could influence how people choose their current environment.

METHODS

Participants

The study population consisted of AWS who were not receiving therapy and AWS who either had just started therapy or were on a waiting list for therapy at the time of the investigation. AWS in therapy (the T group) were recruited from 14 stuttering and/or speech and language therapy centers throughout the Netherlands and from a family system therapy program for persons who stutter. AWS not receiving treatment (the NT group) were recruited informally, by asking relatives and acquaintances of the researchers to invite individuals who stutter and who were currently not in treatment to participate in the study. In addition, a Dutch social networking website for persons who stutter (Hyves-stuttering) was used for recruitment of this group.

Data collection

All data were collected between February and November, 2008. Study questionnaires were distributed by mail. All participants received a small gift for their participation. Besides the outcome measures listed below, all participants were asked to complete a socio-demographic checklist.

Measurement

Symptom status

Self-assessment scale of speech (SA scale)

The primary instrument to assess symptom status was a self-assessment scale of speech (SA scale) [142]. Participants were asked to rate their speech on a scale ranging from 1 (very poor) to 10 (very good). Only the endpoints of the scale were defined (see Figure A3.1 in the Appendix). No normative score is available for this SA scale. Instead, the instrument is criterion-referenced in relation to the Dutch standards (e.g., for school performances) with 1 being the worst and 10 being the best score. A SA score of 6 can be interpreted as speech being sufficiently good.

Speech satisfaction scale

In addition to the SA scale, participants rated their speech satisfaction on a Likert scale with five response categories ranging from "not at all satisfied" to "very satisfied" (see Figure A3.1 in the Appendix). This speech satisfaction scale takes into account more explicitly that symptom status is influenced by intra-individual characteristics.

Overall Assessment of the Speaker's Experience of Stuttering for adults (OASES)

Finally, the OASES [86] was used to assess symptom status. The OASES is a validated questionnaire which evaluates the experience of the stuttering disorder from the perspective of the AWS. It consists of four parts, each of which examines different aspects of the stuttering disorder: (I) general perspectives about stuttering, (II) affective, behavioral and cognitive reactions to stuttering, (III) functional communication difficulties and (IV) impact of stuttering on the speaker's quality of life. Impact ratings scores can be calculated for each individual section and for all sections in total and provide an indication of the impact of stuttering on various aspects of the speaker's life. Although it is emphasized that the impact ratings are not exchangeable with stuttering severity ratings, they may provide an indication of the severity of stuttering [86]. Yaruss and Quesal [86] also presented normative scores. Impact scores between 20.0 and 29.9 refer to mild stuttering, scores of 30.0-44.9 to mild-to-moderate stuttering, 45.0-59.9 to moderate, 60.0-74.9 to moderate-to-severe and 75.0-100 to severe stuttering [86]. For this study, the OASES was translated into Dutch, using the well-established method of forward-translation and back- translation. While sections III and IV of the OASES include outcomes pertaining to functioning and overall QOL, in this study the instrument is classified as a symptom measure, because the OASES has a condition-specific focus; it does not tap *all* aspects of functioning and QOL.

Functioning status and QOL

Health Utility Index 3 (HUI3), EuroQoL EQ-5D and EQ-VAS

General functioning and QOL were measured simultaneously using two widely applied generic instruments that measure functioning and provide a QOL value for the health states that could be described by the instrument: The HUI3 [111] and EQ-5D [113]. Both instruments generate a descriptive health profile of a person's functioning in society on generic, basic domains of life (i.e., physical, mental and social domains). As such, the results of these descriptive systems display a profile of functioning. In addition, a population-weighted health index (or *value*) is produced, based on the descriptive system. This value reflects the general population's perception of the desirability of a health status. In other words, it represents how good or how bad a health state is according to the general population. Values range from -.59 (worst imaginable health state) to 1 (full health) for the EQ-5D [150], and from -.36 to 1 for the HUI3 [111]. The values represent overall QOL scores.

Both the HUI3 and EQ-5D were included because their responsiveness to stuttering has not yet been explored. Although these instruments *conceptualize* health and QOL similarly, the health concept is *operationalized* differently so that differences in responsiveness may be expected. The EQ-5D consists of five domains: mobility, self-care, usual activities, pain/discomfort and anxiety/depression, with three response levels for each domain. The

HUI3 incorporates eight domains: vision, hearing, speech, ambulation, dexterity, emotion, cognition and pain and discomfort (not specified), with five to six response levels for each domain. While, in general, the two descriptive systems lead to similar conclusions about QOL, differences between instruments on QOL values in AWS could be expected. This is because the HUI3 explicitly deals with QOL problems related to speech, while for the EQ-5D inferences about the impact of stuttering on QOL could only be inferred from reduced functioning in other domains like for example anxiety/depression. Although the EQ-5D is currently the most used instrument of the two, the HUI3 is considered preferable in studies focusing on vision, speech or hearing, since these domains are included in the HUI3 and not in the EQ-5D [151]. For stuttering, however, evidence as to the performance of these instruments is lacking.

The EQ-5D was administered in conjunction with the *EQ-VAS* [113]. The EQ-VAS is a visual analogue scale (similar to a thermometer) for recording an individual's rating for his or her current health state. Valuations range from 0 (worst imaginable health) to 100 (best imaginable health). A relevant distinction between the EQ-VAS versus the EQ-5D and HUI3 is that the EQ-VAS values QOL from the perspective of the respondent himself instead of the general population.

Characteristics of the individual: coping
Coping Inventory for Stressful Situations (CISS)
Coping style was measured using the Dutch version of the CISS [149]. Like other coping instruments, the CISS explores an individual's ability to cope with problems by measuring the extent to which that individual applies the various coping styles generally available [147]. Three coping styles are identified by the CISS: task-oriented (T) coping style, emotion-oriented (E) coping style, and avoidance-oriented (A) coping style. The distinction between task and emotion-oriented coping is generally accepted. Task-oriented coping is aimed at actively managing the stressful situation itself, while emotion-oriented coping is aimed at thinking or feeling in a different way about the stressful situation and so reducing the negative emotional consequences [152]. Both types of coping are important, if used properly [152]. The avoidance-oriented coping style refers to the actions aimed at avoiding or withdrawing from the stressor or the feelings that are evoked by the stressor (e.g., daydreaming about other things or meeting friends) [149].

Each coping style is assessed according to 16 items, making a total of 48 items. For each item, respondents indicate on a Likert scale ranging from 1 (not at all) to 5 (very much) to what extent they apply a certain coping strategy during a stressful situation. An example item for T-coping in the CISS is "Think about the event and learn from my mistakes", an example item for E-coping is "Blame myself for being too emotional about the situation"

and an example item for A-coping is "Take some time off and get away from the situation". Studies of the construct validity of the CISS have shown that CISS-T reflects an active and adaptive coping strategy, while CISS-E reflects a more negative way of dealing with emotions. CISS-A could be considered as an active coping strategy [149]. Raw scores can be transformed into normative scores, which are available for the working population and students. Normative scores for the Dutch CISS are classified into seven categories, ranging from 1 (very low use of coping style) to 7 (very high use of coping style). Internal consistency and validity of the CISS is reported to be satisfactory [149].

Analysis

SPSS software version 15.0 (SPSS Inc., Chicago, IL, USA) was used for all the statistical analyses. Categorical variables were described by tabulations and percentages and tested for differences between the T and NT group by Chi-square tests. Continuous variables were described by means and standard deviations and tested for group differences by independent samples T-tests (2-tailed). Cohen's d was used to interpret effect sizes. Cohen defined effect sizes as "small, d= .2; medium = .5 and large = .8" [117].

The primary research question in this study ("To what extent is overall QOL impaired in AWS?") was addressed by calculating the scores on the HUI3, EQ-5D and EQ-VAS. In addition, QOL scores were calculated for and compared between stuttering severity groups. For this analysis, the following stuttering severity categories were created: SA scale: mild = score 8-10; moderate = score 5-7; severe = score 1-4; and for the speech satisfaction scale: mild = score 5 (reflecting high speech satisfaction), moderate = score 2-4 (reflecting low to normal speech satisfaction), severe = score 1 (reflecting very low speech satisfaction). For the OASES, the original normative categories were applied.

To answer the second research question ("What are the determinants of QOL in AWS?") correlation analyses were conducted to explore the relationships between self-assessment of speech, speech satisfaction, OASES Total impact score, coping style, functioning and QOL. Speech satisfaction and coping style utilized Spearman rank correlations; all other comparisons utilized Pearson product-moment correlations. Multiple linear regression analyses were performed to explore the associations of the explanatory variables with QOL. Three regression analyses were run, with respectively the HUI3, EQ-5D and EQ-VAS as dependent variables. In the first step, the SA score (representing symptom status) was entered. To determine the effect of self-assessment of stuttering on QOL with and without the influence of demographic variables, the second step included adding demographic variables (age, gender, education level, marital status), which were entered all at once. In step three, the coping scores for CISS-E, CISS-T and CISS-A were entered. Lastly, the grouping variable (T-NT) was entered. The adjusted R^2 value reflects how well the model

fits the data. In addition to the correlation and regression analyses, comparison of the T and NT group scores provided insight into the determinants of QOL in AWS.

RESULTS

Characteristics of participants and response rate

A total of 91 AWS participated in this study: 38 AWS in the NT group and 53 AWS in the T group (Table 3.1). Significant group differences were found for age (t(89) = 2.390, p = .019) and gender (χ^2(1)= 4.670, p = .031). The response rate for the NT group contacted informally was 92%. In addition, four people responded to the appeal on the Hyves-stuttering website. Two of them were added to the NT group. The other two people were currently in treatment, and consequently added to the T group. Twenty-nine participants in the T group had just started conventional stuttering therapy (mean number of sessions 2.1; S.D. 1.9); six were on a waiting list. Ten people had just entered a family system therapy program for persons who stutter and eight were still on a waiting list for this program.

Symptom status

SA scores and speech satisfaction scores

The results for the SA scale and the speech satisfaction scale are displayed in Table 3.2. For the total group, the mean SA score of 6 corresponds to a rating of speech being sufficiently good. The mean satisfaction score was close to the category "neither satisfied nor dissatisfied." There was a significant group difference (T versus NT) for both the SA scores, t(89) = 3.235, p = .002, effect size = .27, and the speech satisfaction scores, t(89) = 4.136, p < .001, effect size = .43.

OASES Impact scores

Table 3.2 also presents the OASES Impact scores. The mean Total Impact score of 48.4 in the total group represent moderate stuttering. The range of 25.6 to 74.4 indicates that the study population did not contain people with severe stuttering according to the OASES. The OASES Total Impact scores differed significantly between the T and NT group, t89) = -3.728, p < .001, effect size = .80. Significant differences were also found for the individual OASES sections, with higher Impact scores for the T group: Section I: t(89) = -2.380, p = .019, effect size = .51; Section II: t(89) = -3.044, p = .003, effect size = .64; Section III: t(88) = -3.580, p = .001, effect size = .76; Section IV: t(89) = -3.382 , p = .001, effect size = .73.

TABLE 3.1 Demographics

	Total group (N=91)	T group (N=53)	NT group (N=38)	p-value
	N (%)	N (%)	N (%)	
Gender*				.031
Male	63 (69.2)	32 (60.4)	31 (81.6)	
Female	28 (30.8)	21 (39.6)	7 (18.4)	
Age in years*	36 (14.68)	33 (12.43)	40 (16.59)	.019
Educational level				.900
Low	8 (8.8)	5 (9.5)	3 (7.9)	
Middle	28 (30.8)	17 (32.1)	11 (28.9)	
High	55 (60.4)	31 (58.4)	24 (63.2)	
Marital status				.172
Single / divorced	46 (50.6)	30 (56.6)	16 (42.1)	
Married	45 (49.4)	23 (43.4)	22 (57.9)	
Job status				.191
Paid work	60 (65.9)	33 (62.3)	27 (71.1)	
Student	20 (22.0)	15 (28.3)	5 (13.2)	
Other	11 (12.1)	5 (9.4)	6 (15.7)	
Stuttering ever diagnosed by a SLP[a]				.872
Yes	74 (81.3)	44 (83.0)	30 (78.9)	
No	14 (15.4)	8 (15.1)	6 (15.8)	
Unknown	3 (3.3)	1 (1.9)	2 (5.3)	
Onset of stuttering				.066
Onset before 7 years	69 (75.8)	37 (69.8)	32 (84.2)	
Onset ≥ 7 years	22 (24.2)	16 (30.2)	6 (15.8)	
Age of onset in years if onset ≥ 7 years	10.8 (5.34)	11.7 (6.10)	8.7 (1.40)	.307

[a] SLP: Speech-language pathologist
* p < .05 (difference T-NT group, 2-tailed)

TABLE 3.2 Speech characteristics

	Total group (N=91)	T group (N=53)	NT group (N=38)	p-value
	Mean (SD)	Mean (SD)	Mean (SD)	
SA score**	6.0 (1.48)	5.6 (1.44)	6.6 (1.35)	.002
Speech satisfaction score**	2.9 (0.97)	2.5 (0.89)	3.3 (0.90)	.000
OASES Total Impact score**	48.4 (10.88)	51.8 (10.39)	43.7 (9.86)	.000
Impact score Section I*	58.0 (9.68)	60.0 (9.23)	55.2 (9.71)	.019
Impact score Section II**	51.9 (13.64)	55.4 (12.68)	47.0 (13.56)	.003
Impact score Section III**	45.6 (12.32)	49.3 (11.91)	40.5 (11.10)	.001
Impact score Section IV**	39.3 (13.35)	43.1 (12.98)	34.0 (12.10)	.001

* p < .05 (difference T-NT group, 2-tailed)
** p < .01 (difference T-NT group, 2-tailed)

Functioning status and QOL

Descriptive dimensions of functioning

Table 3.3 displays the health profiles for the EQ-5D and the HUI3 by means of frequencies of AWS reporting no problems on the dimensions. The distribution on the speech domain of the HUI3 differed significantly between groups ($\chi^2(1)= 7.595$, $p = .006$), with the T group reporting more problems. For the domains pain/discomfort and anxiety/depression of the EQ-5D and vision, emotion, cognition and pain and discomfort of the HUI3 no significant group differences were established. For the domains mobility, self-care and usual activities of the EQ-5D and hearing, ambulation and dexterity of the HUI3 the number of people in the "problems" cells was too small to allow statistical analyses.

TABLE 3.3 Functioning profiles and quality of life scores

HUI3 dimension	Total group (N=91) N (%) reporting no problems	T group (N=53) N (%) reporting no problems	NT group (N=38) N (%) reporting no problems	*p*-value
Vision	48 (52.7)	30 (56.6)	18 (47.4)	.384
Hearing	88 (96.7)	51 (96.2)	37 (97.4)	a
Speech**	65 (71.4)	32 (60.4)	33 (86.6)	.006
Ambulation	88 (96.7)	52 (98.1)	36 (94.7)	a
Dexterity	88 (96.7)	51 (96.2)	37 (97.4)	a
Emotion	38 (41.8)	21 (39.6)	17 (44.7)	.626
Cognition	68 (74.7)	39 (73.6)	29 (76.3)	.768
Pain and discomfort	53 (58.2)	28 (52.8)	25 (65.8)	.216
EQ-5D dimension	N (%) reporting no problems	N (%) reporting no problems	N (%) reporting no problems	
Mobility	88 (96.7)	52 (98.1)	36 (94.7)	a
Self-care	90 (98.9)	52 (98.1)	38 (100)	a
Usual activities	86 (94.5)	48 (90.6)	38 (100)	a
Pain/discomfort	75 (82.4)	42 (79.2)	33 (86.8)	.348
Anxiety/depression	73 (80.2)	43 (81.1)	30 (78.9)	.796
Overall QOL score	Mean (SD)	Mean (SD)	Mean (SD)	
HUI3	.85 (.16)	.84 (.19)	.88 (.12)	.355
EQ-5D	.93 (.12)	.92 (.14)	.94 (.10)	.520
EQ-VAS**	83.2 (11.9)	80.4 (12.9)	86.9 (9.3)	.007

[a] Chi-square tests could not be performed because the number of people in the "problems" cell was too small
** $p < .01$ (difference T-NT group, 2-tailed)

QOL values

Table 3.3 also displays the overall QOL scores for the HUI3, EQ-5D and EQ-VAS. There were no significant differences between the T and NT group for the HUI3 and EQ-5D. However, both groups differed significantly on the EQ-VAS: $t(89) = 2.772$, $p = .007$, effect size = .81, with a lower score for the T group. Two people in the T group had a very low HUI3 score (.09 and .17), indicating a very low QOL. Removing these outliers did not result in a change in the mean scores for HUI3, EQ-5D and EQ-VAS.

QOL scores differentiated by stuttering severity levels

To explore differences in QOL scores due to stuttering severity, the QOL scores for the total group, differentiated by stuttering severity level, are displayed in Table 3.4. Compared with perfect health (valued at 1), the HUI3 and EQ-VAS scores show a reduction in QOL for adults with mild stuttering (reflected as a high score on the SA scale, a high satisfaction score and a mild OASES Impact score). Furthermore, Table 3.4 shows that QOL reduces with increasing stuttering severity level, irrespective of how stuttering severity was quantified. The HUI3 shows a larger reduction of QOL than the EQ-5D and EQ-VAS.

TABLE 3.4 Quality of life scores by stuttering severity level

	HUI3 score Mean (SD)	EQ-5D score Mean (SD)	EQ-VAS score Mean (SD)
SA score			
Mild (score 8, 9)[a] $n = 11$	.91 (.13)	.96 (.09)	85.9 (10.17)
Moderate (score 5, 6, 7) $n = 65$	.88 (.13)	.93 (.13)	83.1 (11.54)
Severe (score 2, 3, 4)[a] $n = 15$	.73 (.24)	.88 (.14)	81.4 (14.94)
Speech satisfaction score			
Mild (score 5) $n = 4$	.95 (.07)	1.00 (.00)	96.3 (4.11)
Moderate (score 2, 3, 4) $n = 72$	.86 (.16)	.93 (.12)	82.6 (11.76)
Severe (score 1) $n = 5$	.73 (.24)	.88 (.17)	82.0 (14.47)
OASES Total impact rating			
Mild $n = 5$	.96 (.07)	1.0 (.00)	92.0 (7.29)
Mild-to-moderate $n = 30$	.92 (.08)	.96 (.09)	84.3 (12.60)
Moderate $n = 45$	.83 (.15)	.91 (.14)	82.3 (10.31)
Moderate-to-severe $n = 11$	.74 (.28)	.88 (.15)	79.6 (16.51)

[a] No respondents rated their speech with a score of 1 or 10

Characteristics of the individual: coping scores

The internal consistency of the CISS in this study was satisfactory (Cronbach's alpha ranging from .78 to .90). The mean transformed coping scores for the total group for CISS-T were 3.3 (SD = 1.67), for CISS-E 3.9 (SD = 1.65), and for CISS-A 2.4 (SD = 1.37). The mean scores for CISS-E coping are close to average, while the other coping style scores appear to be below average. There were no significant differences between the T and NT group.

Association between symptom status, coping, functioning and QOL

Exploration of the relationships between SA score, speech satisfaction score, OASES Total Impact score, coping and QOL (Table 3.5) revealed that all stuttering symptom measures correlated significantly with each other. In addition, all QOL measures correlated significantly with one or more stuttering symptom measures, with a lower QOL score reflecting more severe stuttering. Overall, the strongest correlations were established for the HUI3, which correlated significantly with all three subjective stuttering measures. The mean EQ-5D QOL score was related to the mean SA score and OASES Total Impact score, but not to the mean speech satisfaction score. The EQ-VAS score only correlated significantly with the OASES Total Impact score. CISS-E was negatively associated with speech satisfaction and the OASES Total Impact score, and was the single coping style significantly related to all QOL measures. CISS-A was positively associated with speech satisfaction, while CISS-T did not correlate significantly with any of the subjective stuttering measures.

TABLE 3.5 Correlations between SA score, speech satisfaction score, OASES Total impact score, coping and overall quality of life

	SA score	Speech satisfaction score	OASES Total Impact score	CISS-T	CISS-E	CISS-V	HUI3 score	EQ-5D score
SA score								
Speech satisfaction score	0.724**							
OASES Total Impact score	-0.701**	-0.638**						
CISS-T	-0.053	0.034	-0.074					
CISS-E	-0.197	-0.251*	0.483**	-0.041				
CISS-V	0.193	0.253*	-0.083	0.240*	0.090			
HUI3 score	0.365**	0.357**	-0.483**	0.159	-0.395**	0.012		
EQ-5D score	0.206*	0.194	-0.336**	0.030	-0.367**	-0.210	0.713**	
EQ-VAS score	0.058	0.137	-0.218*	0.063	-0.382**	-0.029	0.548**	0.451**

* $p < .05$
** $p < .01$ (2-tailed)

Correlations between symptoms and relevant subscales of the HUI3 and EQ-5D (representing functioning) are shown in Table 3.6. The speech and emotion domains of the HUI3 correlated significantly with all three stuttering measures. The cognition and pain domains of the HUI3 and the domains daily activities and anxiety/depression of the EQ-5D all significantly correlated with one stuttering symptom measure.

TABLE 3.6 Correlations between SA score, speech satisfaction score, OASES Total impact score and domains of functioning

	HUI3 speech	HUI3 emotion	HUI3 cognition	HUI3 pain and discomfort	EQ-5D daily activities	EQ-5D anxiety/ depression
SA score	.327**	.274**			-.207 *	
Speech satisfaction score	.294**	.324**				
OASES Total impact score	-.307**	-.384**	-.324**	-.254*		.346**

Note. Only significant correlations are displayed
* *p*< .05
** *p* < .01 (2-tailed)

Regression analysis

The regression model which was used to simultaneously evaluate the effect of each determinant on QOL (Table 3.7) explained 36% of the variation in HUI3 scores (adjusted R^2 full model). Significant independent explanatory variables were SA score (*p* = .006), age (*p* = .003), gender (*p* = .001), marital status (*p* = .045), CISS-T score (*p* = .023) and CISS-E score (*p* = .000). Group identification (group ID) did not contribute to the variation in HUI3 score. The same regression analyses with the EQ-5D score as dependent variable showed only a significant effect of CISS-E (*p* = .000, total adj. R^2 = .186). Regression analyses run with the EQ-VAS as dependent variable showed, in addition to CISS-E, also a significant effect of age (*p* =.001), gender (*p* = .000), marital status (*p* = .005) and group ID (*p* = .000). The total adjusted R^2 was .312.

DISCUSSION

The objectives of the present study were (1) to investigate to what extent QOL is impaired in AWS and (2) to identify determinants of QOL in AWS. The latter was pursued by exploring relationships between stuttering severity, coping, functioning and QOL and by testing for differences in variable scores in two subgroups: the NT group and the T group. The results of this study show that stuttering severity affects overall QOL considerably. HUI3 derived QOL values were .91 for mild stuttering and .73 for severe stuttering. AWS who had just begun or were about to begin therapy rated their stuttering as more severe and recorded more problems on the HUI3 speech domain than AWS who were not in therapy.

TABLE 3.7 Multiple regression analysis for HUI3, EQ-5D and EQ-VAS

	HUI3				EQ-5D				EQ-VAS			
	adjusted R²	delta R²	unstandardized B	p-value	adjusted R²	delta R²	unstandardized B	p-value	adjusted R²	delta R²	unstandardized B	p-value
(Constant)			.674	.000			.826	.000			107.14	.000
Step 1	.123	.133			.032	.043			-.008	.003		
SA score			.032	.006**			.009	.329			-1.119	.185
Step 2	.190	.103			.012	.024			-.002	.050		
Age			-.004	.003**			.000	.406			-.320	.001**
Gender			.117	.001**			.028	.330			9.386	.000**
Educational level			.004	.648			.001	.863			-.625	.326
Marital status			.071	.045*			.039	.183			7.323	.005**
Step 3	.358	.181			.195	.200			.169	.190		
CISS-T			.004	.023*			.002	.163			.210	.056
CISS-E			-.006	.000**			-.005	.000**			-.491	.000**
CISS-V			-.002	.283			-.002	.234			-.098	.437
Step 4	.361	.010			.186	.001			.312	.137		
Group ID			-.037	.257			-.009	.747			-10.036	.000**

* $p < .05$
** $p < .01$

However, the results with respect to the differences in overall QOL between the T and NT group varied. While differences in overall QOL were not significant according to the HUI3 and the EQ-5D, according to the EQ-VAS they were. The effect size was .81, which can be considered as large [117]. The correlation analysis between stuttering severity and domains of functioning in the total group showed that a higher stuttering severity was mainly associated with limitations in the domains of speech and emotion. Lastly, regression analysis showed that the relationship between stuttering severity and overall QOL was influenced by task-oriented and emotion-oriented coping style.

With regard to the extent to which QOL in AWS is affected, our study could not confirm that the impact of severe stuttering on overall QOL was as great as suggested by Bramlett et al. [137]. QOL values for severe stuttering in our study ranged from .73 to .88, while Bramlett et al. found QOL values between .44 and .81 for severe stuttering [137]. There could be two reasons for this difference. Firstly, this might be related to the somewhat wider range of stuttering severity in the study by Bramlett et al. [137]. Although a substantial number of participants in the current study had low scores on the SA-scale, which represents severe stuttering, none of the participants was classified as severe stuttering by the OASES. Secondly, the difference in QOL values may be due to differences in the way QOL values were obtained. Bramlett et al. derived their QOL values by *direct* valuation of vignettes describing stuttering: AWNS rated hypothetical states of stuttering and their own health state [137]. In the current study, QOL was *indirectly* assessed by using the HUI3 and EQ-5D. In this way QOL values (from the general public) were derived by applying a mathematical algorithm to the health states that were described by the AWS. These health states were generic, that is, they had no specific reference to stuttering. Therefore, the indirect instruments applied in the current study might not have been responsive enough to stuttering, resulting in an upward bias. In other words, the impact of stuttering on QOL might actually be greater than found in our study. Alternatively, it could be hypothesized that the absence of anchor points referring to other conditions worse than severe stuttering led to a downward bias in the direct assessment approach by Bramlett et al. [137]. This is known as contextual bias [153]. The two studies have no single measure in common to explore whether the negative impact on QOL has been underestimated in our study or overestimated in the Bramlett et al. study [137], or both.

In the current study, as in the study of Bramlett et al. [137], substantial differences in QOL values were established using different instruments. Comparing the three QOL measurements for the most severe stuttering state, the impairment on the HUI3 was greater than on the EQ-5D and EQ-VAS. This difference may be explained by inclusion of the speech domain in the HUI3, which improves its responsiveness to stuttering. This might also clarify why the HUI3 measurement showed QOL impairment for the mildest forms of

stuttering, but EQ-5D measurement showed relatively little or none. Ceiling effects for the EQ-5D, as reported in other relatively healthy populations [154,155], may have contributed to a limited responsiveness of this instrument in AWS. Accordingly, the EQ-5D might have overestimated QOL, although the alternative hypothesis, that the HUI3 has underestimated QOL, cannot easily be abandoned. By inclusion of speech as a domain, the emphasis on the speech problems may be larger than their impact on QOL warrants.

In theory, EQ-VAS outcomes could help to identify whether QOL was underestimated by the HUI3 or overestimated by the EQ-5D, since the EQ-VAS measures QOL directly and not via its impact on basic domains of functioning. Therefore, the VAS scale is not prone to possible misrepresentation of QOL, which could occur if the HUI3 and the EQ-5D do not include all the relevant domains. In addition, the EQ-VAS values QOL from the perspective of the respondent himself instead of the general population. However, neither hypothesis could be supported, since the results indicate that the EQ-VAS was less responsive than *both* the EQ-5D and HUI3 for changes at the symptom level. An *end of scale* bias might have limited the responsiveness of the EQ-VAS. Subjects tend to avoid using scale ends [88,156], which implies that the QOL effect of mild health problems is difficult to measure on a VAS scale. Support for this hypothesis is found in the result that EQ-VAS scores were limited to a smaller range of the scale than HUI3 and EQ-5D scores. Thus, unfortunately, the EQ-VAS does not provide the key to whether the EQ-5D overestimated QOL, or the HUI3 underestimated it.

Our findings that stuttering affects functioning in a negative way are in line with the results of other studies (e.g., [41,85,132]). The domains that significantly correlate to stuttering severity in our study correspond to a great extent with the domains affected in the Craig et al. study [41], that is mainly social and psychological dimensions. An interesting finding of the current study is the positive correlation between stuttering severity as measured by the OASES and the pain domain of the HUI3. This result may reflect the broad definition of the HUI3 pain domain, which covers pain *and discomfort*. Alternatively, AWS reporting physical pain, especially in the breast region, when asked what they feel in their body when they speak, stutter or try to avoid stuttering, is a quite common response in the clinical experience of the third author. Besides, it may be hypothesized that stuttering affects physical well-being because of higher stress levels associated with the experience of social anxiety [130]. There is evidence for a common neural basis for regulating social pain and physical pain [157]. As a result, the physical pain threshold can be triggered by social pain.

The regression analyses into the relationships between stuttering severity, coping and overall QOL identified coping as a mediating factor in QOL in AWS, in addition to stuttering severity and demographic variables. The results of the HUI3 regression analysis suggested

that both stuttering severity and coping style can be directly related to QOL in equal measure. Two types of coping were associated with QOL. Higher scores on the CISS-E (emotion-oriented coping) were correlated with lower QOL. While it is known that dealing with emotions in a constructive way positively influences the adjustment to a chronic disease [147], higher CISS-E scores reflect a more negative way of dealing with emotions (e.g., denial, mental or behavioral distance, brooding), presumably resulting in a greater psychological impact and a lower QOL [149]. The regression analysis also revealed that higher task-oriented coping scores were associated with better QOL, reflecting that task-orientation is an active and adaptive way of coping which influences QOL in a positive way [158]. QOL might be maximized by individuals who apply the various strategies flexibly depending on the circumstances that they have to deal with [152].

The differences in the results between the therapy group and non-therapy group in this study provide further insight into the underlying mechanisms of QOL in AWS. The groups differed significantly in stuttering severity, in score on the speech domain of the HUI3 and in overall QOL as assessed by the EQ-VAS. There were no group differences in coping scores. The regression analysis with the EQ-VAS as dependent variable was the only analysis that revealed group ID as a significant predictor of overall QOL. These results suggest that AWS who seek treatment do this because they desire symptom relief, and not because they are poor at coping.

Elements in our study design that might evoke questions about the external validity are related to the choice of including a T and NT group of AWS and to the use of self-assessed measures to establish stuttering severity. The NT group was included because we wanted to cover the maximum range of QOL values in the group of AWS and hypothesized that QOL might be higher in AWS not seeking treatment and/or that relationships between stuttering, coping and QOL might differ between groups. The representativeness of the NT group cannot be established, due to the lack of detailed information about the Dutch AWS population not receiving treatment. Furthermore, the results show that there are between group differences, namely a lower stuttering severity and a better subjective QOL, as measured with the EQ-VAS, for the NT group. The difference in stuttering severity was also reflected in a better HUI3 speech QOL value for the NT group. These results imply that outcomes obtained in clinical populations cannot simply be generalized to the population of AWS as a whole and vice versa. With regard to the applied speech measures, we are confident that self-identification of stuttering in the NT group and self-assessed stuttering severity has not negatively affected the external validity, because 81% of the AWS in the study reported having been previously diagnosed as stuttering by a professional. Furthermore, Huinck and Rietveld [142] showed that correlations between a self-assessment scale of speech satisfaction and measures which reflect overt stuttering

behavior are relatively strong, indicating a high validity of a simple and cost-effective speech rating scale. This suggests that our study results would provide a valid estimation of QOL in all AWS.

Our study presents evidence that stuttering in adults is a serious problem affecting health. A broadly based outcome measure such as QOL could provide a means of evaluating the impact of stuttering on daily life. QOL measures could therefore be applied in therapy evaluation studies, or in evaluating the relationship between the cost and benefit of stuttering interventions. Furthermore, the relevance of coping for QOL in AWS, which was demonstrated in this study, shows that a good understanding of the determinants of QOL is essential to develop rational and cost-effective treatments: *"The development of treatment strategies requires not only that we identify the key factors that combine to determine function and quality of life, but also that we understand their relative importance and the degree to which they can be altered or modified"* ([79] p. 63). Our study is a first step in exploring the determinants of QOL in relation to stuttering. The effect of coping on the relationship between stuttering severity and QOL which was established in this study suggests that addressing coping style could be a useful component in the process of diagnosing and selecting treatment approaches for AWS. Using a coping instrument during the assessment phase indicates how an individual copes with stressful situations in daily life. If an AWS is using an inadequate coping pattern, therapeutic goals could be identified which would enable the AWS to change his personal coping style to deal more effectively with stressors that provoke stuttering or the stuttering behavior itself, thereby reducing its negative impact on QOL. For instance, if a client displays relatively high scores on the emotion-oriented coping scale and low task-oriented coping scores, treatment goals might be focused on learning task-oriented coping strategies and becoming less dependent on emotional ways of dealing with stress. This idea is supported by Hayhow et al. [84] who showed that AWS have the desire to get help in managing their stuttering and in developing coping strategies. We would therefore recommend that more studies be done on coping in relation to stuttering, such as the ones recently reported by Plexico and colleagues [83,143-145].

In conclusion, by using generic QOL measures, it was shown that the health condition of moderate to severe stuttering substantially reduces the QOL in AWS as compared to the perfect health state. This result, and the significant relationship between stuttering severity, coping style and QOL, highlights the need for further research in order to clarify the conceptualization of QOL in relation to stuttering, as a foundation for the further development of effective therapies for the disorder of stuttering.

ACKNOWLEDGMENTS

APPENDIX

1. If you had to score your own speech (range 1-10), how would you score it?

    <u>Circle a score</u>

1	2	3	4	5	6	7	8	9	10

    1= very bad

    10= very good

2. How satisfied are you with your speech?

   Mark the corresponding box with a cross.

Not satisfied at all	Not satisfied	Satisfied nor dissatisfied	Satisfied	Very satisfied
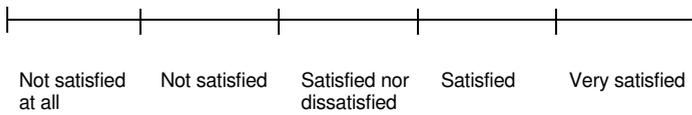
FIGURE A3.1 Self-assessment scale of speech

Psychometric evaluation of the Dutch translation of the Overall Assessment of the Speaker's Experience of Stuttering for adults (OASES-A-D)

With Matthijs Versteegh
Scott Yaruss

ABSTRACT

The Overall Assessment of the Speaker's Experience of Stuttering for adults (OASES-A) [86,159] is a patient-reported outcome measure that was designed to provide a comprehensive assessment of "the experience of the stuttering disorder from the perspective of individuals who stutter" [86]. This paper reports on the translation process and evaluates the psychometric performance of a Dutch version of the OASES-A. Translation of the OASES-A into Dutch followed a standard forward and backward translation process. The Dutch OASES-A (OASES-A-D) was then administered to 138 adults who stutter. A subset of 91 respondents also evaluated their speech on a 10-point Likert scale. For another subset of 45 respondents, a clinician-based stuttering severity rating on a 5-point Likert scale was available. Thirty-two of the respondents also completed the Dutch S-24 scale [160]. The OASES-A-D showed acceptable item properties. No ceiling effects were observed. For 30 out of 100 items, most of which were in Section IV (Quality of Life), floor effects were observed. Cronbach's α coefficients for all sections and subsections surpassed the 0.70 criterion of good internal consistency and reliability. Concurrent validity was moderate to high. Construct validity was confirmed by distinct scores on the OASES-A-D for groups with different levels of stuttering severity as rated by the speakers themselves or by clinicians. These results suggest that the OASES-A-D is a reliable and valid measure that can be used to assess the impact of stuttering on Dutch adults who stutter.

INTRODUCTION

Research in recent decades has shown that stuttering is often associated with negative impact on various aspects of a speaker's life (e.g., [41,42,85,91]). This has led to greater awareness among many researchers and clinicians of the need to adopt broad-based measures that reflect the broader stuttering *disorder* (i.e., the difficulties a person may experience as a result of producing stuttering behaviors, including negative impact on quality of life and subjective well-being), in decision-making, clinical practice, and research (e.g., [9,75,78,86,161]. One measure that was designed for comprehensively assessing the stuttering disorder is the *Overall Assessment of the Speaker's Experience of Stuttering* (OASES) [86,159]. This questionnaire evaluates "the experience of the stuttering disorder from the perspective of individuals who stutter" ([86], p. 90). The design of the OASES was based on the World Health Organization's *International Classification of Functioning, Disability, and Health* (ICF) [80]. There are three versions of the OASES: The OASES-A was designed for adults, ages 18 and above; the OASES-T [162] was designed for teenagers, ages 13-17; and the OASES-S [163] was designed for school-age children, ages 7-12.

Empirical data have provided preliminary support for the reliability and validity of the OASES-A, based on samples collected in the United States [86]. However, analyses conducted to date have not thoroughly examined several aspects of the psychometric properties of the instrument, one of which is convergent validity [75]. Further, data from individuals residing in locations other than the United States have only recently become available (e.g., [93,164-167]. The adult version of the OASES has been translated into Spanish [159] and, at the time of this writing, there are ongoing efforts to translate the various versions of the OASES into approximately 15 other languages worldwide [168]. Key aspects of the translation process involve validation of the translation and evaluation of the psychometric data that result from administration of the translated version to native speakers of the target languages.

Among researchers and clinicians in the Netherlands, a desire exists to have a well-functioning Dutch patient-reported outcome measure in order to be able to assess those aspects of the stuttering disorder that are directly relevant to the lives of people who stutter. To fulfill this need, we translated the English OASES-A into Dutch. In the present study, we describe the translation process and evaluate the psychometric performance of the Dutch version of the OASES-A. We aim to contribute to the evidence base of the performance of the OASES-A in general, and the Dutch translation in particular.

METHODS

The OASES-A questionnaire was first published in 2006 [86] based on preliminary research that had been conducted over the prior 10 years (e.g., [169]). Below, we describe the characteristics of the original instrument, the translation process of the Dutch version and the psychometric evaluation.

OASES-A

The OASES-A is a 100-item, self-report questionnaire that aims to measure the experience of the stuttering disorder from the perspective of adults who stutter. It consists of four sections, each of which examines different aspects of the stuttering disorder: (I) general perspectives about stuttering (20 items); (II) affective, behavioral and cognitive reactions to stuttering (30 items); (III) functional communication difficulties (25 items) and (IV) impact of stuttering on the speaker's quality of life (25 items). Responses are rated on a Likert scale with response choices ranging from 1 to 5. Higher scores indicate a greater impact of the disorder. Impact rating scores can be calculated for each individual section and for all sections in total. These scores provide an indication of the degree of negative impact experienced by a speaker as a result of stuttering. As a self-report measure, the OASES-A is designed to supplement clinician-based measures of observable stuttering severity. Although it is emphasized that the impact ratings are not exchangeable with stuttering severity ratings, they may provide an indication of the overall severity of the speaker's experience of stuttering [86].

In this paper, the Impact scores for the Dutch version of the OASES-A were calculated in accordance with the first version of the OASES-A [86], except where indicated otherwise. Scoring for the 2006 version of the OASES-A involved three steps. First, the number of points the respondent indicated was calculated for each section. Second, the total number of items completed by the respondent was computed and multiplied by 5 (since each item is based on a 5-point scale) to obtain the total number of possible points in each section. Third, the number of points was divided by the number of possible points and multiplied by 100. Impact scores were categorized as follows: 20.0-29.9 refer to mild impact, 30.0-44.9 to mild-to-moderate impact, 45.0-59.9 to moderate impact, 60.0-74.9 to moderate-to-severe and impact, 75.0-100 to severe impact. The scoring system in the current versions of the OASES-A (beginning with the 2008 version and continuing with the 2010 publications) is different, in that the section and overall impact scores are based on the same 1 to 5 range as the individual item scores. Note that it is possible to convert between the two scoring systems by simply dividing the scores from the 2006 version by 20 to yield scores on the 1 to 5 scale used in the 2008 and 2010 versions. Detailed background information and explanations about the development of the OASES-A can be found in publications by Yaruss and Quesal [86,159].

Translation of the OASES-A

The original published English version of the OASES-A [86] was translated into Dutch following a standard forward and backward translation process [170] to ensure conceptual equivalence and clear and easy understanding of the Dutch version of the OASES-A. Initially, items and response choices in the American version of the OASES-A were translated into Dutch independently by two native Dutch speakers who were fluent in English. Then, a first consensus version was produced from the two forward translations. This Dutch consensus version was back-translated into English independently by two qualified translators who are native English-speakers and fluent in Dutch. The research team, which had requested the translation, then compared the back translations with the original version. Problematic items or response choices were discussed in a meeting by the translators and the research team.

A linguistically and conceptually comparable translation generally requires that careful attention be paid to cultural differences that might lead to different meanings in the target and original language [170,171]. Three items were identified as potentially reflecting conceptual differences between the OASES-A and the Dutch translation. These items were carefully discussed in the meeting and consensus was reached regarding the most appropriate translation. The second consensus version of the Dutch OASES-A (hereafter referred to as the OASES-A-D) was pilot tested in a sample of six individuals who stutter. In keeping with recommendations for creating a valid translation [170], pilot testing was also completed with three individuals who did not stutter to ensure that the comprehensibility of the translation was not limited only to people who already possessed some understanding of the stuttering disorder. Participants were asked to complete the questionnaire and comment on the questions if necessary. As a result of the pilot testing, a missing word was added to question II.B.6. No other problems were detected in terms of item acceptance, comprehensibility, or wording or in the consistency of response patterns. This version was used in all subsequent testing.

Data collection procedures

For the psychometric evaluation of the OASES-A-D, we made use of two existing datasets in which the OASES-A-D had been administered to adults who stutter. All data were collected between February 2008 and April 2009. The first dataset (N=91) originated from a study into the quality of life in adults who stutter (hereafter referred to as the "QOL study"). The QOL study included both people who were not receiving therapy and people who had just registered for therapy at the time of the investigation. Demographic characteristics (gender, age, educational level, marital status and job status), OASES-A-D data, and a self-assessment score of speech (SA scale score; [172]) were available from that study. The SA scale was applied to evaluate the participant's perception of his or her stuttering severity.

Participants were asked to rate their speech on a scale ranging from 1 (very poor) to 10 (very good). Only the endpoints of the scale were defined. Further details of the QOL study can be found in Koedoot et al. [91].

The second dataset (N=51) originated from stuttering therapists working in clinics throughout the Netherlands. The therapists asked adults who stutter who had registered for or who were involved in therapy to complete the OASES-A-D and the Dutch S-24 Modification of the Andrews and Cutler [72] adaptation of Erickson's scale of communication attitudes (S-24) [160]. The S-24 is a self-completed questionnaire which measures the communication attitudes of persons who stutter. Besides the two self-reported questionnaires, the therapists also rated the stuttering severity of their clients on a 5-point Likert scale with the following categories: 1 = mild, 2 = mild-moderate, 3 = moderate, 4 = moderate-severe, and 5 = severe stuttering. When rating severity, the therapists were asked to take into account the speaker's total experience of the disorder, including cognitive, emotional, motor and social aspects. In the rest of the paper this scale is referred to as the Clinical Assessment (CA) scale. Since all therapists had many years of experience in diagnosing and treating people who stutter and because they are accustomed to classifying stuttering severity of clients in terms of mild, moderate and severe stuttering, the CA scale was considered an appropriate measure of stuttering severity. The S-24 data were available for 32 participants and the CA scale data for 45 participants.

In total, 142 people who stutter completed the OASES-A-D (91 participants in the QOL study and 51 participants recruited by therapists). The data from four participants were excluded in the present study because they were less than 18 years of age. Thus, this study was based on the responses of 138 participants. Demographic characteristics of these participants are presented in Table 4.1. More men than women participated in our study. The male: female ratio of 2.7:1 is generally comparable with ratios presented in literature (e.g., [7]). Compared to data of Statistics Netherlands (CBS, http://www.cbs.nl/en-GB/menu/home/default.htm) a relatively high proportion (that is, 50%) of the participants had received higher education. There were no respondents with a minority ethnic background (e.g., Moroccan, Turkish or Surinamese).

Item characteristics
The OASES-A-D item performance characteristics that were studied included item distributions and percentage floor and ceiling effects (i.e., the percentage of respondents scoring at respectively the lowest and highest scale level).

TABLE 4.1 Demographics

	N	Dutch population norms[a]
Gender		
Male	101 (73.2%)	49.5%
Female	37 (26.8%)	50.5%
Age (years)		
Mean (SD)	34.5 (12.8)	40.1
Range	18-74	-
Educational level[b]		
Low	6 (6.8%)	33%
Middle	36 (40.9%)	31%
High	44 (50%)	27%
Missing	2 (2.3%)	9%
Marital status[b]		
Single / divorced	43 (48.9%)	-
Married	45 (51.1%)	-
Job status[*]		
Paid work	60 (68.2%)	-
Student	17 (19.3%)	-
Other	11 (12.5%)	-

[a] Statistics Netherlands, 2009 figures
[b] Only available for participants in the QOL study

Reliability

Internal consistency refers to the extent to which items within each domain are interrelated, thus reflecting the degree to which they measure the same concept. Cronbach's α coefficient is the most widely applied method to assess internal consistency (e.g., [173]). A coefficient of above 0.70 suggests a good internal consistency and reliability [174], however, if α is too high, this may suggest a high level of item redundancy [175]. In addition to the Cronbach's α scores of Sections I to IV, we assessed each subsection of Section II to IV individually, since pooling the scores within a section could inflate Cronbach's α due to the large number of items. The division of Section I (General information) in three subsections was done merely for convenience in scoring the record form; the items are not conceptually related. Therefore, Cronbach's α values were not calculated for these subsections.

Validity

In keeping with the original validation process of the English version of the OASES-A, concurrent validity was evaluated by calculating Spearman correlation coefficients for each section of the OASES-A-D and for the Total Impact score with the Dutch version of the S-24. Based on the results of the publication by Yaruss and Quesal [86], the S-24 scores were expected to have high correlations with the OASES-A-D Impact scores from Section II, and moderate correlations with the other sections. In addition, the correlation between the OASES-A-D, the SA, and the CA scores were used for assessing concurrent validity. A strong correlation was considered to be over .60, a moderate correlation between .30 and .60, and a low correlation below .30 [176].

The method of known-groups comparisons was used to evaluate the construct validity of the OASES-A-D. Known-groups validity is defined as the ability to distinguish between clinically relevant subgroups of respondents. We tested if OASES-A-D Total Impact scores could discriminate between participants with different stuttering severity levels. Severity levels were determined by both self-assessed severity (SA scale score) and clinician-assessed severity (CA scale score). Because of the relatively small sample sizes for some categories of stuttering severity, the following categories of the SA scale were merged to reach a sufficient number of respondents in each category: mild = score 7 - 10; moderate = score 4 - 6; severe = score 1 - 3. For the CA scale, the categories were combined as follows: mild = score 1 - 2; moderate = score 3, severe = score 4 - 5.

We also tested whether the OASES-A-D Total Impact score was dependent on the demographic characteristics age and education. For the variable *age*, a correlation coefficient was calculated. For *educational level*, three groups were compared: low (primary education), middle (secondary education) and high (advanced degree).

Statistical methods

Values are reported as mean +/- 1 SD or as absolute number and percentage. One-way analysis of variance (ANOVA) and Tukey post-hoc tests were employed to evaluate the statistical significance of differences in OASES-A-D Impact scores for groups with different levels of stuttering severity and different educational levels. All correlations were based on non-linear Spearman rank correlations, and a Bonferroni correction was applied to maintain an overall α of .05. Analyses were performed in SPSS version 17.0 (SPSS Inc.).

RESULTS

Stuttering characteristics

Table 4.2 presents the mean scores on the OASES-A-D and the other stuttering measurement instruments (i.e., the SA scale and CA scale) applied in this study. To facilitate comparison of results from the OASES-A-D with the current version of the OASES-A, as well as results obtained from translations of the OASES-A in other languages, Table 4.2 also reports the mean scores in accordance with the 5-point scale scoring system introduced in Yaruss and Quesal [159] and used in all three of the current OASES record forms. All other tables and results in this paper use the scoring system from the original 2006 publication, as described above in Section 2.1.

Item characteristics

All but 15 of the 100 items of the OASES-A-D exhibited ranges from the minimum possible score of 1 to the maximum possible score of 5. The mean score across items ranged from 1.32 to 3.74 (SD ranging from 0.65 to 1.46). No ceiling effects (defined as > 30 % of patients having the maximum score of 5) were observed. Floor effects were observed for 30 out of 100 items, most notably in Section IV (Quality of Life) with 14 items. Section IV.D (which measures the impact of stuttering on job and education) showed floor effects for four out of five items, indicating that respondents experienced relatively little negative impact from stuttering in these settings. Section IV.E (which measures the impact of stuttering on overall well-being) showed floor effects for six out of eight items.

TABLE 4.2 Stuttering characteristics

Stuttering instrument	Mean, SD based on original scoring procedures described in Yaruss & Quesal [86]	Mean, SD based on revised scoring procedures described in Yaruss and Quesal [159]
OASES-A-D Impact scores		
Section I	56.8 (10.37)	2.84 (0.52)
Section II	52.2 (12.66)	2.61 (0.63)
Section III	46.5 (11.86)	2.32 (0.59)
Section IV	40.1 (13.21)	2.00 (0.66)
Total	48.7 (10.45)	2.44 (0.52)
SA score[a]	6.11 (1.41)	
CA score[b]	3.09 (1.12)	

[a] Only available for participants in the QOL study
[b] Only available for participants recruited by therapists

Reliability

Cronbach's α scores for Sections I through IV, as well as for the subsections of Section II to IV, of the OASES-A-D are presented in Table 4.3. Cronbach's α scores for the four sections were between 0.84 and 0.96. The subsections showed Cronbach's α values between 0.78 (Section III.C) and 0.92 (Section IV.E).

TABLE 4.3 Cronbach α of the OASES-A-D sections

OASES-A section	Number of items	Cronbach's α
I	20	0.84
II	30	0.93
II.A	10	0.9
II.B	10	0.82
II.C	10	0.81
III	25	0.94
III.A	10	0.84
III.B	5	0.86
III.C	5	0.78
III.D	5	0.8
IV	25	0.96
IV.A	3	0.8
IV.B	4	0.84
IV.C	5	0.89
IV.D	5	0.9
IV.E	8	0.92

Validity

The Total OASES-A-D Impact score, as well as the Impact scores on the four sections, correlated significantly with the S-24, SA and CA scale scores (Table 4.4). For the S-24 and the CA scale, the lowest correlations were established for Section I and the highest for Section IV. For the SA scale, the pattern was reversed, with a slightly lower correlation for Section IV.

Table 4.5 shows that all sections of the OASES-A-D questionnaire were able to discriminate between groups of participants with different stuttering severity levels (according to the SA score or the CA score), with the exception of discriminating between participants with moderate and severe stuttering as assessed by the SA scale.

TABLE 4.4 Correlations (Spearman rho) between OASES-A-D Impact scores and S-24, SA scale and CA scale scores

OASES-A section	S-24 (N=32)	SA scale (N=91)	CA scale (N=45)
Impact score Section I	.587**	-.609**	.357*
Impact score Section II	.641**	-.507**	.561**
Impact score Section III	.761**	-.543**	.494**
Impact score Section IV	.854**	-.516**	.572**
Total Impact score	.838**	-.615**	.594**

* $p < .05$ level (2-tailed)
** $p < .01$ (2-tailed)

TABLE 4.5 Mean OASES-A-D Total Impact scores for participants with mild, moderate and severe stuttering according to the SA scale and CA scale, standard deviation (SD) and *p*-value of ANOVA-analysis for differences of means

Stuttering severity level (SA scale)				Significance (*p*)		
Mild (N=38)	Moderate (N=46)	Severe (N=4)	F-ratio	Mild vs. moderate stuttering	Mild vs. severe stuttering	Moderate vs. severe stuttering
41.7 (9.2)	51.6 (8.5)	58.4 (9.5)	16.336	<.001	.002	.314
Stuttering severity level (CA scale)				**Significance (*p*)**		
Mild (N=13)	Moderate (N=17)	Severe (N=14)	F-ratio	Mild vs. moderate stuttering	Mild vs. severe stuttering	Moderate vs. severe stuttering
43.1 (6.4)	51.4 (10.0)	59.2 (7.9)	12.381	.027	<.001	.037

TABLE 4.6 Correlations (Spearman rho) between OASES-A-D Impact scores and age ($p > .10$)

OASES-A section	Age (N=138)
Impact score Section I	-.039
Impact score Section II	-.055
Impact score Section III	-.173
Impact score Section IV	-.112
Total Impact score	-.111

The OASES-A-D Total Impact score, as well as the Impact scores on the sections I, II and IV, did not correlate significantly with age (see Table 4.6, $p > .10$). There was a very small relationship between the Impact score on Section III and age ($r = -.173$, $p = .04$), but after Bonferroni adjustment for the significance level (1/5 * .05 = .01) this was not significant. No significant differences in impact score were detected based on level of education (see Table 4.7, $p > .10$).

TABLE 4.7 Mean OASES-A-D Impact scores for participants with low, middle and high education, standard deviation (SD) and *p*-value of ANOVA-analysis for differences of means

OASES-A section	Educational level			F-ratio	Significance (*p*)		
	Low (N=6)	Middle (N=36)	High (N=44)		Low vs. middle education	Low vs. high education	Middle vs. high education
Impact score Section I	55.9 (8.6)	57.5 (10.9)	58.5 (8.9)	.222	.925	.821	.906
Impact score Section II	55.3 (6.9)	52.4 (14.3)	49.9 (12.7)	.649	.868	.613	.682
Impact score Section III	49.1 (6.5)	45.7 (12.4)	43.8 (11.8)	.649	.795	.562	.755
Impact score Section IV	39.2 (4.2)	40.8 (13.5)	36.5 (12.2)	1.207	.953	.871	.274
Total Impact score	49.8 (4.3)	49.0 (11.4)	46.7 (9.9)	.602	.982	.768	.587

DISCUSSION

In this article, we have reported on the translation and psychometric characteristics of the Dutch version of the OASES for adults (OASES-A-D). The OASES-A-D showed acceptable item properties, a good internal consistency and moderate-to-high significant correlations with other existing instruments. The translated questionnaire showed no ceiling effects, and the majority of the items exhibited ranges from the lowest possible score of 1 to the highest possible score of 5. For fifteen out of 100 items, the maximum score did not reach 5, which can be explained by the relatively small number of participants in this study with severe stuttering. The mean scores across items ranged from 1.32 to 3.74 (*SD* ranging from 0.65 to 1.46), showing similar variability as that seen in Yaruss and Quesal [86], who found a range of the mean from 1.7 to 3.5 (*SD* 0.75 to 1.6). Floor effects were observed most frequently in Section IV (Quality of Life). This may suggest that the OASES-A-D questionnaire lacks some sensitivity on the lower end of the scale, especially in the sections on job and education (IV.D) and overall well-being (IV.E). However, as our sample included mainly people with mild or moderate stuttering, the item scores probably adequately represent the impact of relatively mild stuttering on these aspects of quality of life. The findings regarding potential floor effects thus need further empirical evaluation.

The reliability of the translated questionnaire was assessed using only internal consistency. All four sections of the OASES-A-D demonstrated strong internal consistency, with Cronbach's α scores greater than 0.90 for Sections II to IV, and a Cronbach's α score of 0.84 for Section I. Scores were thus well above the 0.70 required to support internal consistency [174]. They were also in line with the results on the internal consistency reported by Yaruss and Quesal [86], who found Cronbach's α values between 0.92 and 0.97. Cronbach's α scores for each subsection were also above 0.70.

To assess concurrent validity, correlations between Impact scores and the Dutch S-24, SA and CA scores were calculated. Overall, concurrent validity was moderate to strong. The highest values were obtained for the correlations between the OASES-A-D and the S-24. Correlations between the OASES-A-D sections with the S-24 in our study ranged from .59 to .85. This range was in line with values found for preliminary versions of the OASES-A in the United States, i.e., .68 to .83 [86]. However, in our study, the highest correlation was established for Section IV (Quality of Life) and not, as was anticipated, for Section II (Reactions to Stuttering). The correlations between the Total Impact score and the two different measures of stuttering severity applied in this study (the SA scale, measuring subjective stuttering severity, and the CA scale, measuring the clinician's rating of stuttering severity) were both approximately .60. Since there are fundamental differences between the instruments in the way stuttering is evaluated (i.e., the SA scale measures stuttering severity by means of a self-rating of speech on a 10-point scale, the CA scale represents a clinician-based judgment, and the OASES-A-D comprehensively assesses the participant's experience of the stuttering disorder), these correlations are judged to represent adequate relationships. Finally, age and educational level had little or no influence on the OASES-A-D Impact scores. The lack of a correlation between OASES-A-D scores and chronological age is consistent with prior preliminary reports [177]. These findings support the concurrent validity of the OASES-A-D.

Another way to measure validity is to compare groups known to differ on relevant features (known-group or construct validity). All sections of the OASES-A-D were able to differentiate between groups of participants with different levels of stuttering severity. Only the moderate and severe categories of the SA scale did not show significant differences in mean OASES-A-D Impact score. However, this may be due to the fact that this test was underpowered, since only four participants in this sample reported severe stuttering problems.

Our study has several limitations. First, test-retest reliability of the OASES-A-D was not assessed. Prior research [86,159] has revealed high test-retest reliability for the original English version of the OASES-A, though further research will be needed to determine the test-retest reliability of the Dutch version. Second, not all of the questionnaires that were used in our psychometric analyses were available for all participants. For the participants recruited by therapists, no SA score was reported. Due to the fact that the data for the other participants were extracted from an ongoing QOL study, not all instruments that were relevant for the current study were applied. As a result, the CA scale scores and S24 scores were missing for those participants. Third, to perform a known-group analysis, categories of the SA scale were combined to create three groups (mild-moderate-severe stuttering), since we did not have enough data to perform the analysis with five groups. The same

was done for the CA scale. Even after combining categories, however, the distribution of stuttering severity scores on the SA scale remained skewed, with only four people reporting severe stuttering. In future studies, it would be recommended to include a more balanced sample with respect to stuttering severity. Even with these limitations, however, results support the general conclusion that the Dutch translation of the OASES-A exhibits appropriate psychometric properties.

The current study yielded some results that point to areas for improvement in future revisions of the OASES-A-D in particular and the OASES-A in general. Particularly, Cronbach's α values for Section II, III and IV were above 0.90, indicating that there might be redundant items in these sections. Although it typically requires only 15 or 20 minutes to complete, the OASES-A is a relatively long questionnaire. The potential benefit of this is that it provides detailed information to clinicians about their clients' experience of the stuttering disorder [86,159]. Still, for some clients, the length of the form may cause some concern. To reduce this burden and for reasons of parsimony, a shorter questionnaire targeted particularly for use in research may also be beneficial (though some of the detail inherent in the tool that helps clinicians with treatment planning and goal setting may be diminished). Future research could provide more insight into the possible redundancy of some items. Shortening the questionnaire could be based on several arguments. First, additional analysis may reveal that reducing the number of items with high correlations within a subsection may not reduce the sensitivity of the instrument. Second, item response theory might provide evidence for the redundancy of items and answer categories. A preliminary Rasch analysis [178,179] that we performed suggested that Sections I and II had a better fit to the Rasch model when the answer categories were rescored to a four point scale. Thus, in addition to considering the length of the questionnaire, the number of response categories could be evaluated. Such modifications to the questionnaire are beyond the scope of this paper, as any adaptation would require renewed psychometric testing. Therefore, these and other improvements to the OASES-A remain an interesting avenue of future research.

To conclude, this study provides preliminary results that the Dutch language version of the OASES-A is a reliable and valid instrument for providing a comprehensive assessment of how stuttering affects the lives of individuals who stutter. Findings are relevant both to individuals who are in therapy as well as to those who are not. The fact that translations of the various versions of the OASES are being developed for several languages will, in the future, facilitate the comparability of OASES results in cross-cultural settings. Furthermore, it provides an excellent opportunity for collaborative research between nations.

ACKNOWLEDGMENTS

4

5

Direct versus indirect treatment for preschool children who stutter: the RESTART randomized trial

With Elly Stolk
Toni Rietveld
Marie-Christine Franken

ABSTRACT

Objective Stuttering is a common childhood disorder. There is limited high quality evidence regarding options for best treatment. The aim of the study was to compare the effectiveness of direct treatment with indirect treatment in preschool children who stutter.

Methods In this multicenter randomized controlled trial with an 18 month follow-up, preschool children who stutter who were referred for treatment were randomized to *direct* treatment (Lidcombe Program; n=99) or *indirect* treatment (RESTART-DCM treatment; n=100). Main inclusion criteria were age 3-6 years, ≥3% syllables stuttered (%SS), and time since onset ≥6 months. The primary outcome was the percentage of non-stuttering children at 18 months. Secondary outcomes included stuttering frequency (%SS), stuttering severity ratings by the parents and therapist, severity rating by the child, health-related quality of life, emotional and behavioral problems, and speech attitude.

Results Percentage of non-stuttering children for direct treatment was 76.5% (65/85) versus 71.4% (65/91) for indirect treatment (Odds Ratio (OR), 0.6; 95% CI, 0.1-2.4, p=.42). At 3 months, children treated by direct treatment showed a greater decline in %SS (significant interaction time x therapy: β=-1.89; t(282.82)=-2.807, p= .005). At 18 months, stuttering frequency was 1.2% (SD 2.1) for direct treatment and 1.5% (SD 2.1) for indirect treatment. Direct treatment had slightly better scores on most other secondary outcome measures, but no differences between treatment approaches were significant.

Conclusions Direct treatment decreased stuttering more quickly during the first three months of treatment. At 18 months, however, clinical outcomes for direct and indirect treatment were comparable. These results imply that at 18 months post-treatment onset, both treatments are roughly equal in treating developmental stuttering in ways that surpass expectations of natural recovery. Follow-up data are needed to confirm these findings in the longer term.

INTRODUCTION

Developmental stuttering is a prevalent childhood disorder. The incidence rate is 5 to 11% in preschool years [11,12]. The cause of stuttering is unknown, although recent research indicates that structural and functional brain anomalies underlie the disorder [RW.ERROR - Unable to find reference:3272], with a strong genetic involvement [10,14,17,182]. Many children recover spontaneously; about 63% at 3 years post onset [12,34]. Knowledge of factors that favor the chance for recovery [34,38] can help pediatricians and speech-language pathologists (SLPs) to identify children at risk for chronic stuttering [183]. Nevertheless, the chance for recovery cannot be predicted for an individual child. Since chances for full recovery diminish when stuttering has been present for 15 months [44] and persistent stuttering in adolescents and adults can have a serious mental and social impact [41,91,130], treatment is generally recommended to start before the age of 6 years [11,184]. However, the evidence base for the effectiveness of current therapies for preschool children who stutter is surprisingly weak as well as unbalanced in terms of published reports [66].

For about three decades, many preschool children who stutter around the world have been treated according to an *indirect*, multifactorial treatment approach, like treatment based on the Demands and Capacities Model (DCM) [47,55]. This approach aims to decrease demands set by the environment (e.g., parents are trained to slow down their habitual speech rate) and the child him- or herself (e.g., desensitization for disfluency), and increase the child's capacities for speaking fluently (e.g., accurate and smooth speech motor movements that are age-appropriate) to arrive at a favorable balance between demands and capacities, eventually resulting in fluent speech. Since 2000, an increasing number of children have been treated according to a *direct* operant treatment approach: the Lidcombe Program (LP) for early intervention [48,185]. This direct approach teaches parents to give verbal contingencies after fluent and stuttered speech. With the limited data available at present, the direct LP offers the best evidence-based intervention for preschool children who stutter [66]. However, the long-term effectiveness of this treatment is still unclear [186]. More importantly, comparative effectiveness to current standard treatment has not yet been established; yet child health policy-makers, pediatricians and SLPs need this information to decide upon reimbursement and treatment choice. This is for instance illustrated by a recent proposal of the national speech-language pathology association of Australia (Speech Pathology Australia) to only fund treatment by the LP [187]. Therefore, the aim of the current study was to compare the effectiveness of direct versus indirect stuttering treatment in preschool children during an 18 month follow-up.

METHODS

Study design, participants and setting

This parallel group randomized trial named RESTART (the Rotterdam Evaluation Study of Stuttering Therapy in preschool children- a Randomized Trial) included 199 preschool children who stutter, who were registered at one of the 20 participating speech clinics (including 24 SLPs) throughout the Netherlands. Eligible participants were children (1) aged 3.0-6.3 years, (2) with a stuttering severity rating ≥ 2 ("mild") on an 8-point scale [34] provided by the parent (3) and by the clinician, (4) who stuttered ≥ 3% of syllables and (5) for at least 6 months. The inclusion criterion of at least 3% syllables stuttered (SS) had replaced the original criterion of "at least 3.3% Stuttering Like Disfluencies (SLD)" shortly before the start of the trial. This was based on critics on the SLD measure in literature and on the results of a study into the validity of the SLD measure that we conducted at our center. Exclusion criteria were a diagnosis of an emotional, behavioral, learning or neurological disorder, or a lack of proficiency in Dutch for children or parents. The exclusion criterion of having received treatment for stuttering during the past year was omitted after 5 months, since it was noticed that by excluding these children, the external validity would be restricted. All SLPs were trained and experienced in both treatments. DCM based treatment training is included in the regular clinical education in the Netherlands, and all but one SLP had additionally been trained in the assessment and treatment of children who stutter to become a certified fluency expert recognized by the Dutch association of stuttering therapy (NVST). To ensure a uniform application of DCM based treatment, a treatment manual was developed in collaboration with all participating clinicians prior to the start of the trial. In addition, all SLPs had gone through a three day LP course taught by a LP Consortium trainer and had been certified to provide LP therapy. They had on average 15 years of experience with DCM based treatment (range 7-21 years) and 3.7 years with the LP (range 1.5-7.6 years). Therapists' fidelity to treatment was monitored in 3-monthly intervision meetings, regular telephone contacts with the research team, and by registration forms on the content and amount of treatment filled in by the SLPs and checked by the research team. The intervision meetings were chaired by a LP consortium trainer and a DCM trainer. The trial was approved by the Ethics Committee of the Erasmus MC and registered at isrctn.org (ISRCTN24362190). Written informed consent was obtained from all parents. The trial protocol and supporting CONSORT checklist are available as supporting information: see S1 Checklist and S1 Protocol (online).

Interventions

Direct treatment: The Lidcombe Program

The Lidcombe Program (LP) is a behavioral treatment based on the premise that stuttering is an operant behavior that can be targeted by contingencies. The LP is administered by parents under the direction of a clinician. Children allocated to the LP were treated according to the LP manual [48]. Parents were trained to deliver verbal contingencies in conversations with their child (e.g., "That was smooth" or "Were there any bumpy words?") in a 5:1 ratio for stutter-free and stuttered speech. During the first stage of the program, the parent delivered contingencies during structured conversations of 10-15 minutes once or twice a day. The speech clinic was attended once a week. This continued until stuttering either disappeared or reached an extremely low level (≤1% of syllables stuttered). During the second stage, the use of verbal contingencies as well as the number of clinic visits was gradually reduced, provided that fluency was maintained.

Indirect treatment: The RESTART Demands and Capacities Model based treatment

RESTART Demands and Capacities Model based treatment (RESTART-DCM) is premised on the idea that positive changes in the child's functioning and/or in the environment will lead to a reduction of stuttering. Following the RESTART-DCM manual [188], parents were trained to decrease relevant motoric, linguistic, emotional or cognitive demands, thereby reduce communicative pressure on the child (e.g., parents slowing down their habitual speech rate). If deemed necessary, the child's capacities for fluency were subsequently trained (e.g., improving the child's speech motor movements or word-finding capacity). Parents were required to give their child their undivided attention and practice home assignments 15 minutes a day, for a minimum of 5 days a week. Treatment was gradually reduced if the child showed acceptable speech, parents had mastered implementing a fluency enhancing environment and knew what to do if a relapse occurred.

Randomization and blinding

A minimization software program (MINIM2) [189] was used by the principal investigator (CdeS) to allocate children to one of the treatment arms, according to factors known or thought of to be related to treatment outcome [190]: gender, stuttering severity in the clinic (based on the SSI-3 score) [116], time since onset (TSO; 6-12, 13-18, 19+ months), a first, second, or third degree relative with persistent stuttering (yes, no) and/or a history of recovered stuttering (yes, no), stuttering treatment during the past 12 months (yes, no), and SLP. Three stuttering severity categories were distinguished: (1) mild (SSI-3 score: 10-16); (2) moderate (SSI-3 score: 17-26); severe (SSI-3 score: 27+). For each participant, treatment allocation depended on the characteristics of the children already enrolled [190]. Judges of stuttering frequency were blinded to treatment allocation and measurement moment.

5

Outcome assessment

The primary outcome measure was the percentage of non-stuttering children at18 months, operationalized as ≤1.5% syllables stuttered (SS). This criterion was obtained by applying a conversion ratio of 1.15 to the mean percentage of stuttered word disfluencies in children who do not stutter (1.29%) [191]. Parents were requested to make three audio recordings of 10-15 minutes each in a period of two weeks: one sample of their child speaking to a parent at home, one to a non-family member at home and one to a non-family member away from home [6,192,193].

Secondary outcome measures assessed at baseline, and at 3, 6, 12 and 18 months after start of treatment, were the frequency of stuttering (%SS), a severity rating of stuttering by the parent on an 8-point scale[34], and parents´ valuation of their child´s health-related quality of life on a proxy version of the EuroQoL EQ-VAS [113] with anchor points 0 (worst imaginable health) and 100 (best imaginable health). Secondary outcome measures assessed at baseline and 18 months were the speech attitude of the child (KiddyCAT) [96] and emotional and behavioral problems measured by the Child Behavior Checklist (CBCL) [194]. The latter consists of the scales Internalizing (anxiety, depression, withdrawal, and somatic complaints), Externalizing (aggressive and delinquent behavior), and Total problem behavior [194]. At 18 months both the SLP and the child provided a stuttering severity rating: the SLP on an 8-point scale [34], the child on a 4-point scale where 1= I do not stutter anymore and 4= I stutter a lot.

Eight SLPs not involved in the study were trained to determine the %SS of the samples in real time with sufficient intrajudge reliability, using an electronic, button press counter. To ensure sufficient interjudge reliability, 64% of all samples were scored by at least two raters. Disagreements in ratings were discussed and a third, blinded senior rater was consulted in rare cases where no agreement could be reached (cf. Boberg & Kully [195]).

Statistical analysis

An a priori power calculation to detect a difference of 15% in percentage of non-stuttering children (80 versus 95%) with a power of 80% in a 2-tailed test at a significance level of .05 and allowing a 22% drop-out rate, resulted in a sample size of 98 in each group. Baseline factors were characterized as medians, means and standard deviations for continuous variables and as frequency distributions for categorical variables. Baseline comparisons between treatment groups and between survivors and drop-outs were assessed using χ^2 tests and independent t-tests. Participants were analyzed in the group to which they were randomized.

The effect of treatment on the primary outcome measure was analyzed by χ^2 tests and logistic regression analysis (ENTER method). The regression analysis included the main effect of therapy and the interaction terms therapy*age in years, therapy*stuttering severity (SSI-3 score), and therapy*TSO. Confidence intervals around the obtained percentages of children classified as non-stuttering were calculated according to the method of Wilson [196,197], using a website calculator (http://www.vassarstats.net/prop1.html). In a sensitivity analysis, cut-off scores of 1% SS and 2% SS were applied to further assess the robustness of the primary outcome.

For the secondary outcomes assessed at all measurement moments (%SS, parental rating of stuttering severity, and EQ-VAS) and at baseline and 18 months (KiddyCAT and CBCL), we applied a longitudinal repeated-measures mixed effects model with random intercepts, assuming missing at random. Participant was included as a random predictor; fixed predictors were therapy, and 4 cross-products as interaction terms: time*therapy, and time*therapy*age, severity, and TSO, respectively. An unstructured covariance matrix was assumed for the error as a more plausible autoregressive covariance structure did not provide a better fit. This approach was also used at level 2 of the model. Since the data on %SS did not meet the assumptions needed to calculate CIs for the intraclass correlation coefficient (ICC), interjudge reliability of the speech samples was assessed using Krippendorff's alpha [198] with the option "interval data" for the macro developed by Hayes (2013) [199]. For the outcome %SS, an additional analysis was conducted into the progression in the first 3 months. CBCL outcomes at 18 months were analyzed separately using ANOVA-analysis. Secondary outcome measures only assessed at 18 months (severity ratings by clinician and child) were compared by independent t-tests. For all secondary outcomes, unadjusted and Holm-adjusted [200] p-values are presented, using an overall level of significance of α=.05 (2-sided). The Holm's correction is generally considered a good alternative to the conservative Bonferroni approach [201]. Each p_j is compared to α/(n-j+1); that is: the smallest p_j (j=1) is compared to α/n, the next smallest to α/(n-1) etc.

Treatment intensity was compared by independent t-test, and a χ^2 test was conducted to compare the number of children on treatment at the endpoint of the trial. For analysis of the questionnaires, instructions offered in the manuals were followed. All analyses were carried out in SPSS 20 (Armonk, NY: IBM Corp.).

RESULTS

Participants

Children were enrolled between September 2007 and June 2010. Of 615 children referred for treatment, 416 were not eligible for various reasons (Figure 5.1). In total 199 children met the inclusion criteria. One child was found ineligible after inclusion and therefore excluded from all analyses (Figure 5.1). Baseline characteristics did not differ between treatment groups (Table 5.1). In the LP group 12 children were lost to follow-up as compared to 9 children in the RESTART-DCM group (n=21, 11% drop out rate). Children who were lost to follow-up did not significantly differ on any baseline characteristics (age, gender, ethnicity, educational level of parent, SSI-3 score, %SS, TSO, parental ratings, stuttering in family, prior treatment for stuttering) from children who completed the trial (p-values ranging from .11 to .91). For 191 children, at least one outcome measurement after the start of treatment was available.

TABLE 5.1 Baseline characteristics of participants by treatment group

Characteristic	Lidcombe Program (n=98)[a]	RESTART-DCM (n=100)[a]
Age in months, median; mean (SD)	51.0; 51.5 (9.5)	52.0; 54.1 (11.1)
Age in years		
3[b]	41 (41.8)	37 (37.0)
4	39 (39.8)	31 (31.0)
5-6	18 (18.4)	32 (32.0)
Male	68 (69.4)	70 (70.0)
SSI-3 score		
mild[c]	32 (32.7)	31 (31.0)
moderate	47 (48.0)	51 (51.0)
severe	19 (19.4)	18 (18.0)
% SS, median; mean (SD)[d]	4.9; 6.2 (4.4)	4.0; 5.3 (4.3)
Time since onset		
6-12 months	43 (43.9)	45 (45.0)
13-18 months	25 (25.5)	22 (22.0)
19+ months	30 (30.6)	33 (33.0)
Family history of persistency[e]	45 (45.9)	45 (45.0)
Family history of recovery[e]	27 (27.6)	25 (25.0)
Prior treatment for stuttering	8 (8.2)	6 (6.0)

[a] Data are shown as No. (%) unless specified otherwise
[b] One child in the LP group was 2.11 years at time of inclusion
[c] Children with a stuttering frequency < 3% in the therapy setting but ≥ 3% in the home setting were included in the group "mild stuttering" (n=26)
[d] For one child in the RESTART-DCM group %SS on baseline was not available
[e] For one child in the LP group information on family history of stuttering was not available
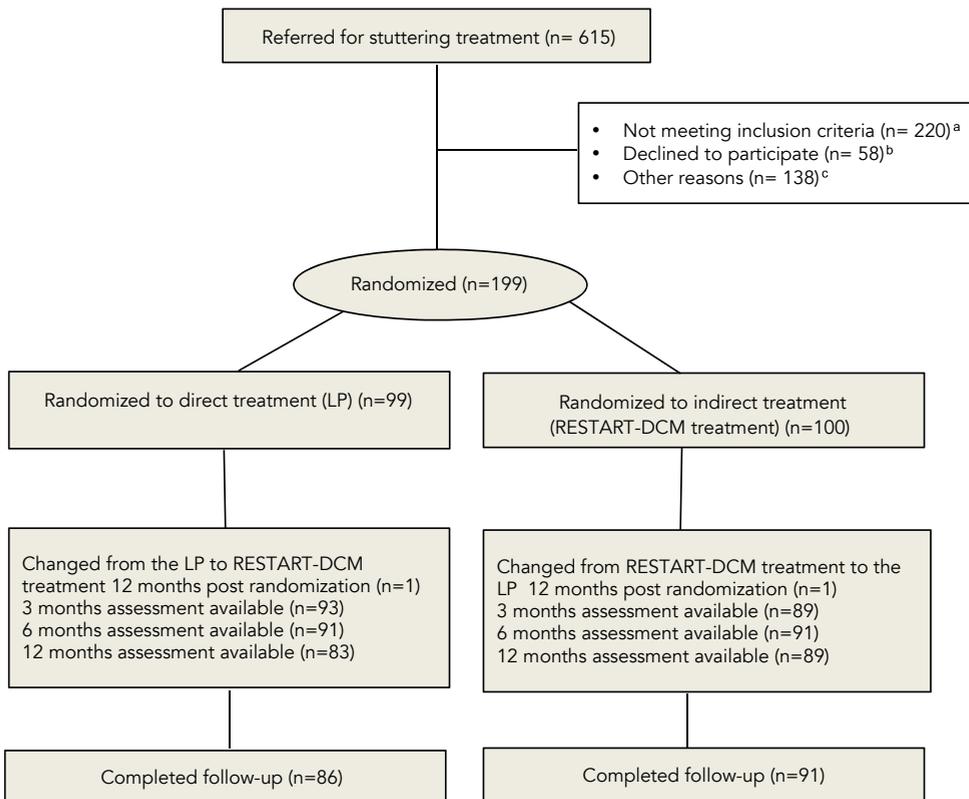
FIGURE 5.1 RESTART trial flow diagram

[a] Borderline stuttering (n=97); Stuttering frequency decreased during assessment phase (n=79); Stuttering <6 months (n=26); Lack of proficiency in Dutch for child or parents (n=18)

[b] Expectations of high demands (n=31); Preference for either treatment (n=27)

[c] Start treatment for other speech/language disorders (n=41); Preference for treatment center closer to parent's home (n=21); Problematic or complex home situation (n=43); Parents did not show up (n=5); Other reasons (n=28)

Speech samples

The mean number of available audio samples for a child at a measurement moment was 2.9 (range 1-6). At least 85% of all samples had a length of ≥300 syllables. The mean intrajudge reliability [202] of measurement of %SS was 83%. Krippendorff's alpha for samples with 2 ratings at baseline and after 3, 6, 12 and 18 months, respectively, was 0.849, 0.896, 0.817, 0.795, and 0.830; all significant, with significance obtained by bootstrapping. All scores represent good reliability [203].

TABLE 5.2 Primary and secondary outcomes at baseline and 18-month follow up

Outcome measure	Baseline			18 months			Parameter	Estimate (95% CI)	z	Unadjusted p-value	Adjusted p-value
	Number of participants	LP	RESTART-DCM	Number of participants	LP	RESTART-DCM					
% recovery	197	-	-	176	76.5	71.4					
Therapy type							OR	0.6 (0.1; 2.4)	-	.42	-
Therapy type x Age(1)[a]							OR	2.8 (0.7; 11.5)	-	.16	-
Therapy type x Age(2)[a]							OR	1.2 (0.4; 3.8)	-	.75	-
Therapy type x Severity(1)[b]							OR	0.6 (0.1; 2.7)	-	.50	-
Therapy type x Severity(2)[b]							OR	0.8 (0.2; 3.0)	-	.71	-
Therapy type x TSO(1)[c]							OR	1.0 (0.3; 3.4)	-	.98	-
Therapy type x TSO(2)[c]							OR	4.6 (0.8; 27.2)	-	.09	-
% SS	197	6.2 (4.4)	5.3 (4.3)	176	1.2 (2.1)	1.5 (2.1)					
Therapy type							β	0.62 (-0.65; 1.89)	0.96	.34	-
Time							β	-0.76 (-1.21; -0.31)	-3.30	.001	.002
Time x Therapy type							β	-0.51 (-0.86; -0.16)	-2.90	.004	.008
Time x Therapy type x Age							β	0.04 (-0.02; 0.10)	1.40	.16	.32
Time x Therapy type x Severity							β	0.04 (-0.01; 0.10)	1.57	.12	.13
Time x Therapy type x TSO							β	0.05 (0.002; 0.11)	2.04	.04	0.13
Parental severity rating	189	4.4 (1.0)	4.3 (1.0)	176	1.0 (1.4)	1.4 (1.5)					
Therapy type							β	0.13 (-0.25; 0.51)	0.68	.50	-
Time							β	-0.67 (-0.85; -0.50)	-7.61	<.001	<.001
Time x Therapy type							β	-0.38 (-0.55; -0.22)	-4.62	<.001	<.001
Time x Therapy type x Age							β	0.07 (0.03; 0.10)	3.66	<.001	<.001
Time x Therapy type x Severity							β	0.04 (0.00; 0.07)	2.15	.03	.10
Time x Therapy type x TSO							β	0.00 (-0.02; 0.04)	0.51	.61	1

TABLE 5.2 Primary and secondary outcomes at baseline and 18-month follow up (Continued)

Outcome measure	Baseline			18 months			Parameter	Estimate (95% CI)	z	Unadjusted p-value	Adjusted p-value
	Number of participants	LP	RESTART-DCM	Number of participants	LP	RESTART-DCM					
EQ-VAS	182	88.0 (10.2)	88.4 (10.1)	168	91.5 (9.7)	90.5 (10.2)					
Therapy type							β	-0.09 (-3.22; 3.04)	-0.06	.96	-
Time							β	0.18 (-0.99; 1.37)	0.31	.76	.76
Time x Therapy type							β	0.35 (-0.79; 1.49)	0.60	.55	.55
Time x Therapy type x Age							β	0.04 (-0.15; 0.35)	0.78	.44	.44
Time x Therapy type x Severity							β	0.10 (-0.47; 0.01)	-1.86	.06	.13
Time x Therapy type x TSO							β	-0.23 (-0.18; 0.27)	0.37	.71	1
CBCL Internal score	193	10.4 (7.9)	7.4 (5.9)	173	5.5 (5.2)	4.2 (4.5)					
Therapy type							β	4.80 (1.21; 8.39)	2.63	.009	.02
Time							β	-1.17 (-4.16; 1.82)	-0.77	.44	-
Time x Therapy type							β	-0.77 (-3.00; 1.46)	-0.68	.50	-
Time x Therapy type x Age							β	-0.29 (-0.63; 0.05)	-0.82	.10	.29
Time x Therapy type x Severity							β	0.14 (-0.19; 0.46)	-1.66	.41	-
Time x Therapy type x TSO							β	-0.13 (-0.44; 0.18)	-0.83	.41	-
CBCL External score	193	13.6 (7.4)	11.2 (7.6)	173	7.1 (5.8)	6.2 (5.7)					
Therapy type							β	3.93 (0.37; 7.49)	2.18	.03	.03
Time							β	-2.85 (-5.68; -0.02)	-1.99	.05	-
Time x Therapy type							β	0.97 (-1.25; 3.19)	0.86	.39	-
Time x Therapy type x Age							β	-0.68 (-1.06; -0.30)	-3.51	.001	.004
Time x Therapy type x Severity							β	-0.03 (-0.40; 0.34)	-0.17	.86	-
Time x Therapy type x TSO							β	-0.07 (-0.41; 0.28)	-0.39	.70	-

5

TABLE 5.2 Primary and secondary outcomes at baseline and 18-month follow up (Continued)

Outcome measure	Baseline			18 months			Parameter	Estimate (95% CI)	z	Unadjusted p-value	Adjusted p-value
	Number of participants	LP	RESTART-DCM	Number of participants	LP	RESTART-DCM					
CBCL Total problem score	193	36.2 (20.6)	27.9 (17.6)	173	21.8 (15.4)	18.6 (13.8)					
Therapy type							β	13.40 (3.75; 23.03)	2.74	.007	.02
Time							β	-3.52 (-11.63; 4.59)	-0.86	.39	-
Time x Therapy type							β	-3.73 (-9.97; 2.50)	-1.18	.24	-
Time x Therapy type x Age							β	-0.50 (-1.31; 0.75)	-0.54	.59	.59
Time x Therapy type x Severity							β	-0.28 (-0.73; 1.25)	0.52	.60	-
Time x Therapy type x TSO							β	0.26 (-1.43; 0.43)	-1.06	.29	-
KiddyCAT[d]	182	3.6 (2.5)	3.9 (2.9)	116	1.2 (1.5)	2.0 (2.1)					
Therapy type							β	0.15 (-1.32; 1.63)	0.20	.84	-
Time							β	-1.35 (-2.77; 0.07)	-1.87	.06	-
Time x Therapy type							β	-0.88 (-1.93; 0.17)	-1.65	.10	.40
Time x Therapy type x Age							β	0.09 (-0.07; 0.25)	1.13	.26	.52
Time x Therapy type x Severity							β	0.04 (-0.11; 0.18)	0.49	.62	-
Severity rating by clinician	NA	NA	NA	168	1.1 (1.4)	1.4 (1.4)					
Therapy type							β	0.00 (0.00; 0.00)	-	.93	-
Therapy type x Age							β	0.08 (0.01; 0.13)	-	.01	.01
Therapy type x Severity							β	0.04 (0.00; 0.09)	-	.14	-
Therapy type x TSO							β	0.02 (0.00; 0.05)	-	.50	-
Severity rating by child	NA	NA	NA	168	1.4 (0.5)	1.4 (0.5)					
Therapy type							β	0.00 (0.00; 0.03)	-	.49	-
Therapy type x Age							β	0.09 (0.01; 0.14)	-	.006	.01
Therapy type x Severity							β	0.04 (0.00; 0.09)	-	.14	-
Therapy type x TSO							β	0.01 (0.00; 0.01)	-	.88	-

a Age(1) refers to age 4 years; age(2) refers to age 5-6 years
b SSI(1) refers to moderate stuttering severity; SSI(2) refers to severe stuttering severity
c TSO(1) refers to TSO 13-18 months; TSO(2) refers to TSO 19+ months
d The KiddyCAT was only applicable for preschool children. Therefore, the effect of TSO could not be precisely estimated and TSO was left out in the analysis

Primary outcome

At 18 months, audiotapes were available for 173 children. For 1 child in the LP and 2 in the RESTART-DCM group audiotapes were missing and replaced by videotapes made in the clinic. For 1 child in the LP group, both audio and videotapes were lacking. Thus, the final analysis at 18 months was based on 176 children. In total, 76.5% (65/85; 95%CI: 66.4-84.2) of children in the LP group were classified as non-stuttering at 18 months compared to 71.4% (65/91; 95%CI: 61.4-79.7) of children in the RESTART-DCM group. This difference was statistically non-significant ($\chi^2(1)$=0.579, p=.45). Nor did logistic regression analysis indicate therapy or other factors as significant predictors of being classified as non-stuttering (therapy: OR, 0.6; 95% CI, 0.1-2.4; p= .42; Table 5.2). Applying cut-off criteria of 1% SS and 2% SS did not significantly affect the results.

Secondary outcomes

The results for all secondary outcome measures at baseline and 18 months and the results for the mixed model analyses are presented in Table 5.2. For the outcome %SS, the effect of therapy type was non-significant. However, a significant interaction between time and therapy type was detected (adjusted p= .008), indicating that the %SS differed for therapy groups at different time points. The effect of time was also significant (adjusted p= .002), indicating that in both treatment groups the average %SS decreased significantly over time. Effect sizes were small (Table 5.2).

Figure 5.2 shows that in both groups most improvement in %SS occurred in the first 3 months of therapy. For this interval, an effect of therapy type was found (β=2.30; t(217.38)=2.10, p= .04), as well as a significant interaction between time and therapy type (β=-1.89; t(282.82)=-2.81, p= .005). Compared to the RESTART-DCM group, the LP group had a slightly higher mean %SS at baseline and showed a greater decline, resulting in a lower %SS at 3 months. Significant interactions with very small effect sizes were also present between time, therapy type, and stuttering severity (β=0.25; t(173.94)=2.51, adjusted p= .01) and time, therapy type, and TSO (β=-0.21; t(172.85)=2.40, adjusted p= .02) (Figure 5.3).

For the outcome parental rating of stuttering severity, a significant effect of time (adjusted p< .001) as well as a significant interaction between time and therapy type (adjusted p< .001) was detected. Figure 5.2 shows a slightly greater decline in scores for the LP group over the period of 18 months. The interaction between time, therapy type and age was significant (adjusted p< .001) but showed a very small effect size (Table 5.2). For the outcomes EQ-VAS and KiddyCAT, no significant effect of therapy type or any other factor was found (Table 5.2; Figure 5.2).
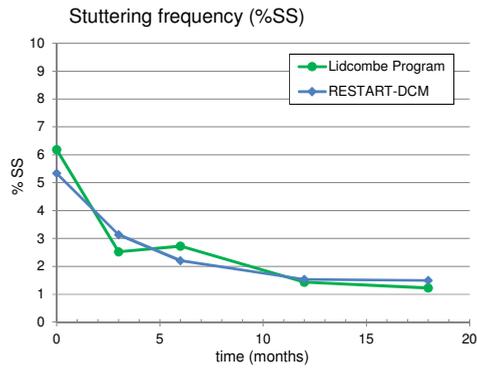
Stuttering frequency (%SS)

No. of children

LP	98	92	85	83	85
RESTART-DCM	99	88	89	86	91

Parental rating of stuttering severity
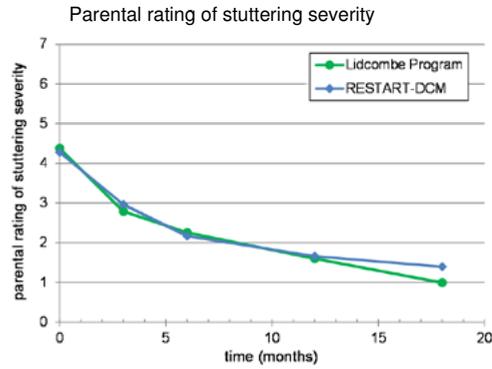
No. of children

LP	98	92	85	83	85
RESTART-DCM	99	88	89	86	91

Note. Scale ranges from 0 (no stuttering), 1 (borderline stuttering), to 7 (very severe stuttering) [11]

EQ-VAS Quality of Life score
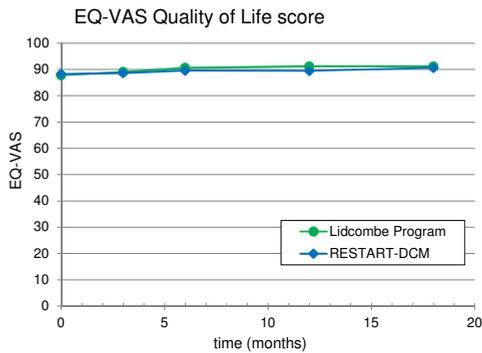
No. of children

LP	90	91	78	79	80
RESTART-DCM	92	86	81	83	88

FIGURE 5.2 Change in three secondary outcome measures during 18-month follow up

For all CBCL scale scores, the factor therapy type was significant (Table 5.2), but this effect was attributable to significantly higher scores for the LP group at baseline. At 18 months, no significant differences were found (Internal scale: $F_{(1,196)}$=-1.04, unadjusted p= .32, partial eta squared=.006; External scale: $F_{(1,196)}$=1.04, unadjusted p= .31, partial eta squared=.006; Total problem scale: $F_{(1,196)}$=1.12, unadjusted p= .29, partial eta squared=.006). For the CBCL External scale, a significant interaction with a small effect size was established between time, therapy type and age: older children showed a greater decline in score, particularly in the LP group.

For the severity rating by the clinician as well as by the child at 18 months, significant interactions between therapy type and age were established (Clinician: adjusted p= .01; Child: adjusted p=.01). However the small eta-squared values (0.079 and 0.088, respectively) suggest that these differences are negligible.
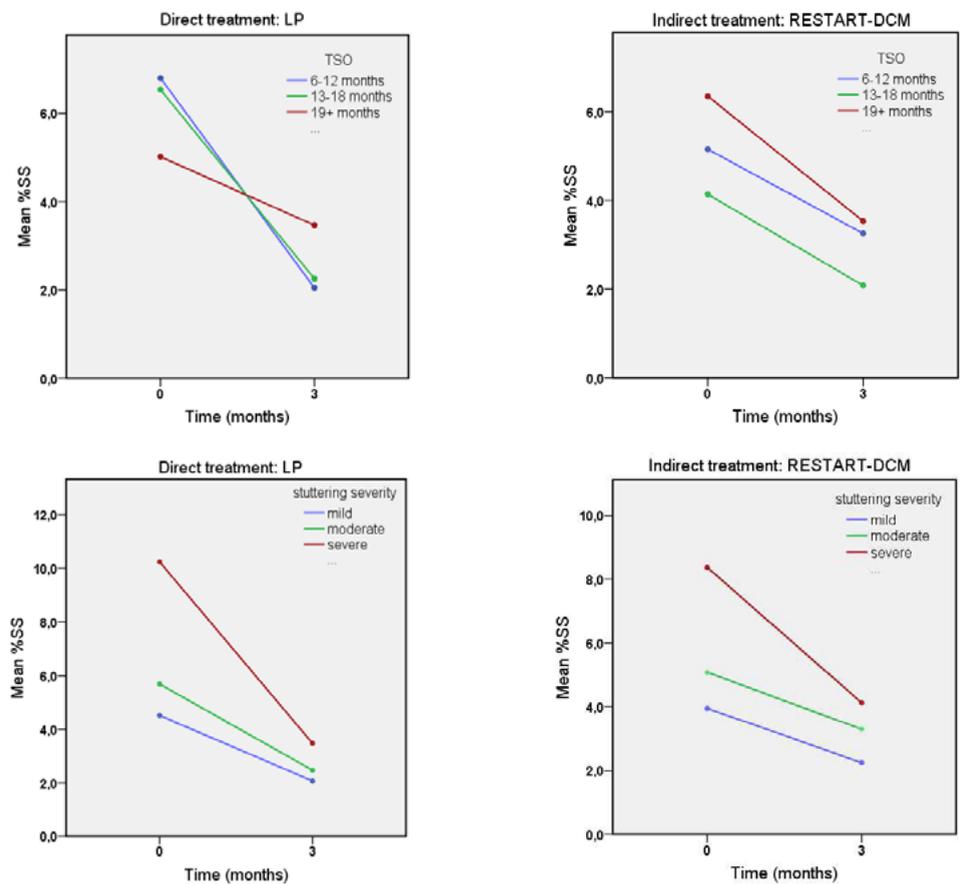
FIGURE 5.3 Change in %SS during first three months

Treatment intensity

The number of treatment sessions and treatment hours did not differ significantly between groups (Table 5.3). At 18 months, 27.6% (27/99) children in the LP group compared to 35.0% (35/100) children in the RESTART-DCM group were still on treatment, a difference that was also not statistically significant ($\chi^2(1)=1.277$, $p=.26$).

TABLE 5.3 Treatment intensity by treatment group

	LP (N=98)	RESTART-DCM (N=97)	*p*-value
Number of treatment sessions, median; mean (SD; SE) [range]	21; 22.2 (11.2; 1.1) [2-51]	17; 19.5 (10.3; 1.0) [2-59]	.08
	LP (N=95)	**RESTART-DCM (N=93)**	*p*-value
Number of treatment hours, median; mean (SD; SE) [range]	18.3; 19.6 (10.9; 1.1) [1.4-51]	15.5; 18.0 (9.7; 1.0) [3.0-55.2]	.20

DISCUSSION

The RESTART-trial found that both direct and indirect treatment for preschool children who stutter reduced stuttering during 18 months of follow-up. The direct approach reduced stuttering frequency more quickly during the first three months of treatment, however, the difference was not significant anymore by 18 months. Most outcome measures were slightly in favor of the direct approach (LP), but the few significant interaction terms were deemed negligible due to their small effect sizes. For most children, stuttering frequency plateaued after three months, while about 30% of children were still on treatment after 18 months.

The direct LP and indirect RESTART-DCM treatment are based on different premises and assumptions regarding mechanisms underlying treatment effect (i.e., delivering verbal contingencies versus balancing demands and capacities for fluent speech, respectively). However, since results for both treatments were comparable, it could be hypothesized that their common components have a larger influence on recovery than their unique components (cf. Imel & Wampold [204]). In psychotherapy and counseling, this is known as the "dodo bird phenomenon" [205]. According to this hypothesis, treatments that are intended to be therapeutic are equally efficacious. Studies suggest that 30-70% of therapy outcome can be attributed to common factors, including good therapeutic relationships [204]. Unfortunately, little is known about the unique mechanisms that lead to change in stuttering behavior in both treatments [206-208]. Common components of the LP and RESTART-DCM treatment may include consideration of maintaining factors, an increase in one-on-one time that parents spend with their child, a boost of encouragement and a reduction of linguistic demands for the child [209], and emotional support for the parents.

Our results do not enable us to distinguish the potential effect of treatment from spontaneous recovery. Spontaneous recovery in the general population at 36 months post onset has been estimated to be 63% or higher [34]. An estimate of the mean time since onset of stuttering at the endpoint in our study is 33 months. Thus, our percentages of children classified as non-stuttering exceed this estimate by about 10% (p=.02; based on statistical test for comparing two proportions from different populations). Furthermore, the chance of spontaneous recovery in our clinical study population is likely to be lower than in the general population, because this chance is known to diminish after 12 to 18 months [34,44] and 56% of children within our study stuttered for at least 12 months.

Strengths of our study are the large sample size with minimal loss to follow-up, the broad range of outcome measures, the large number of measurement moments, and the relatively long follow-up period (double the time in Jones et al. [6]). Participating therapists in the RESTART-study worked in usual-care centers throughout the Netherlands. Thus, the treatments were studied in a variety of regular clinical settings with therapists unconnected to the developers of the therapies [207,210], therefore employing a practical study design ensuring a high external validity. A limitation of our study is that a high number of children appeared ineligible for participation. Results may therefore not be fully generalizable to all preschool children presenting to a clinic with stuttering. Another limitation is that the applied follow-up time is insufficient to decide conclusively whether a child has recovered from stuttering. This requires a period of about 5 years [34,211], to account for the possibility of a relapse. Therefore, we intend to follow-up all children under study.

CONCLUSIONS

At 18 month post-treatment onset, the evidence suggests that both direct and indirect treatment for stuttering can be recommended. However, direct treatment decreased stuttering more quickly during the first three months. Future research investigating the role of client and clinician factors, the effectiveness of a combined direct and indirect approach, and the cost-effectiveness of a limitation of treatment time or frequency may shed further light on the effectiveness of stuttering treatment in preschool children.

ACKNOWLEDGMENTS

6

Economic evaluation of stuttering treatment in preschool children: RESTART randomized trial

With Clazien Bouwmans
Marie-Christine Franken
Elly Stolk

ABSTRACT

Purpose The purpose of this study was to evaluate the incremental cost-effectiveness and cost-utility of the Lidcombe Program (LP) compared with treatment based on the Demands and Capacities Model (RESTART-DCM) for preschool children who stutter.

Methods A cost-effectiveness and cost-utility analysis were carried out alongside a Randomized Clinical Trial (the RESTART-study). In total, 199 children in 20 speech clinics participated. Outcome measures included the number needed to treat, based on the percentage of children who did not stutter at 18 months, and Health-related quality of life (EQ-VAS and HUI3) at 3, 6, 12 and 18 months. Health-related quality of life scores were used to calculate quality adjusted life years (V-QALYs for the EQ-VAS and U-QALYs for the HUI3). Direct and indirect costs were measured by cost questionnaires. Missing data were multiply imputed. Percentages of children who did not stutter in both groups were compared by a chi-square test. Between-group differences in mean QALYs and costs, as well as cost effectiveness and cost-utility ratios, were evaluated by applying bootstrapping techniques.

Results After 18 months, health outcomes were slightly better in the LP group, although only the difference in V-QALYs was statistical significant (0.018; 95% CI: 0.008 to 0.027) with a small effect size (Cohen's d=0.17). Mean costs for the LP group were significantly higher compared to the RESTART-DCM group (€3199 versus €3032), again with a small effect size (Cohen's d=0.14). The incremental cost-effectiveness ratio was €3360 for one additional child who did not stutter with the LP, and the estimated cost-utility ratios were €10413 (extra cost per extra V-QALY) and €18617 (extra cost per extra U-QALY). The results indicated a high probability that the LP is cost-effective compared to RESTART-DCM treatment given a threshold for willingness-to-pay of €20.000 per QALY.

Conclusions Differences in effects and costs between the LP and RESTART-DCM treatment were small. Cost-effectiveness and cost-utility ratios were in favor of the LP. The LP is considered a good alternative to RESTART-DCM treatment in Dutch primary care.

INTRODUCTION

Persistent stuttering can lead to a decreased health-related quality of life (HRQOL) in the psychological, emotional and social domains of functioning [41,91,130], as well as to substantial health care costs [212]. To prevent stuttering becoming persistent, treatment is best initiated in the preschool years. Treatment should preferably result in a high percentage of children recovering from stuttering at acceptable societal costs. Information on costs and effects of available stuttering treatments could help clinicians and policy makers in decisions on therapy choice and reimbursement. Although the last decade has shown an increasing number of studies into the efficacy of stuttering therapy in preschool children, there is a lack of evidence on the cost-effectiveness of available treatments.

Two widely applied treatment approaches for preschool children who stutter are treatment based on the Demands and Capacities Model (DCM) [47,55] and the Lidcombe Program (LP) [185]. In the Netherlands, children are commonly treated according to the former approach. Currently, about 10% of Dutch speech-language pathologists (SLPs) working in private practices are also trained in the LP. The LP is supported by a larger body of evidence than any other treatment [66], but a head-to-head comparison against other types of treatment is presently unavailable. However, it is known that the LP requires a relatively long maintenance phase after fluent speech has been attained. The LP is therefore expected to be more costly than DCM based treatment. The average number of treatment sessions for DCM based treatment has been estimated at 12 sessions [47], while the LP requires almost double. This raises the question whether the presumably higher treatment cost of the LP is compensated by a greater proportion of recovered children, fewer relapses, and better individual speech outcomes, as suggested by Onslow et al. [185].

An economic evaluation can provide insight into the costs and effects of a new health care intervention compared to usual care. All types of economic evaluations assess costs, but health consequences can be measured in different ways [88]. The most common forms of economic evaluation are cost-effectiveness analysis (CEA) and cost-utility analysis (CUA). In a CEA, the health consequences are expressed in terms of natural units (i.e., survival or a desired clinical outcome like recovery), while in a CUA the effects are valued in terms of generic measures of health, such as quality adjusted life years (QALYs) [88]. The comparison of costs and effects of a new intervention with usual care results in an incremental cost-effectiveness or cost-utility ratio. This metric can be used to judge whether the additional effects are large enough to justify the extra costs. To get the most benefit from resources available to society and, accordingly, to guide implementation and reimbursement decisions, an economic evaluation should be conducted from a societal perspective. This implies that all costs and health benefits are included, regardless of to

whom costs are related to or who receives the benefits [88]. In the field of speech-language pathology economic evaluations are scarce, but crucial to provide a basis for decisions on implementation and reimbursement of therapies. The aim of the present study was to determine the incremental cost-effectiveness and cost-utility of the LP compared to DCM based treatment.

MATERIALS AND METHODS

Study design

The economic evaluation was performed alongside a prospective randomized clinical trial in the Netherlands (the RESTART-study) with a time horizon of 18 months. Data was collected between September 2007 and January 2012. A societal perspective was adopted for the economic evaluation. Details of the study design and the interventions have been previously published [213]. The trial was approved by the Ethics Committee of the Erasmus MC, the Netherlands, and has been registered at isrctn.org as ISRCTN24362190.

Participants and setting

Children were included by 24 SLPs in 20 speech clinics throughout the Netherlands. Inclusion criteria were: (1) age between 3.0 and 6.3 years; (2) stuttering confirmed by a rating of stuttering severity on an 8-point scale of at least 2 ("mild") [34] by the parent (3) as well as by the clinician; (4) at least 3% of syllables stuttered (SS); and (5) stuttering for at least six months. Exclusion criteria were: (1) diagnosis of an emotional, behavioral, learning or neurological disorder; and (2) lack of proficiency in Dutch for children or parents. The method of minimization was applied to allocate children to one of the treatment arms [189]. In total, parents of 199 children gave their informed consent and these children were included in the study. This sample size allowed us to detect a 15% difference in percentage of children who did not stutter at 18 months, assuming a power of 80% with an alpha-level of .05.

Interventions

The Lidcombe Program (LP) is a behavioral treatment based on operant methods. Children allocated to the LP were treated according to the "manual for the Lidcombe Program of early stuttering intervention" [48]. The intervention mainly comprises delivering verbal contingencies during conversations when the child is speaking mostly stutter-free (e.g., "That was smooth" or "Were there any bumpy words?"). Parents are trained by the therapist how to deliver the contingencies during weekly clinic visits. Treatment consists of two stages. The median number of clinic visits for stage 1 has been estimated to be 11 [48] to 15 [214]. When stuttering either disappears or reaches a satisfactorily low level, the child

enters stage 2. This maintenance phase comprises at least seven treatment sessions [48], and has been recently estimated to include 8 to 12 sessions [214]. Combining the most recent estimates, the mean total treatment time is 22 to 27 sessions.

DCM based treatment is premised on the assumption that stuttering develops when a child lacks the capacities to speak as fluently as the environment demands. Therefore, treatment aims to achieve a favorable balance between environmental demands (e.g., parents slowing down their habitual speech rate) and demands by the child him- or herself and the child's capacities for speaking fluently (e.g., improving the child's speech motor movements or his word finding capacity). Children allocated to DCM based treatment were treated according to the RESTART-DCM treatment manual [188]. This manual was designed before the onset of the trial, in order to standardize the DCM based treatment approach (hereafter referred to as: RESTART-DCM treatment). At the onset of the treatment, the child attends the clinic every week. The intensity of treatment is gradually reduced if stuttering frequency decreases. The mean number of sessions is estimated to be 12, however with a high variability [47]. In the pilot study that was accomplished before the RESTART-study, the mean number of treatment sessions of DCM based treatment of one hour each in the first three months was 11, and three-quarter of children were still on treatment at the end of the three months [56].

Data collection

At baseline, relevant demographic characteristics (e.g., age, gender, and risk factors for persistency and recovery of stuttering) were assessed. At baseline and 3, 6, 12 and 18 months post-treatment onset, parents were asked by the SLP to make three audio recordings of daily conversations of the child outside the clinic and fill in questionnaires on health outcomes, resource use and costs. Questionnaires on resource use and costs asked about the last three months. To account for costs in the period between six and nine months post-treatment onset, an extra questionnaire was filled in by parents at nine months.

Health outcomes

The primary outcome used for the CEA was the number needed to treat (NNT), based on the absolute percentage of children in both groups who did not stutter at 18 months. The NNT stands for the average number of patients who need to be treated for one patient to benefit compared with a control. The absolute percentages could not directly be applied in the calculation of the cost-effectiveness ratio, as the implication of the costs to obtain one percent more non-stuttering children with either treatment depends on the population size (i.e., €500 for one percent more children who do not stutter in a population of 1000 children differs from, for example, €500 for one percent more children who do not stutter in a population of 2000 children). Thus, for a meaningful application of the percentages of

non-stuttering children in the CEA, these percentages were converted into the NNT. In the RESTART-trial, the NNT stands for the number of children who need to be treated with the LP in order to have one extra child defined as non-stuttering at 18 months post-treatment onset, compared to RESTART-DCM treatment. The NNT is the inverse of the absolute difference in rate of non-stuttering and is computed as $1 / (p_{LP} - p_{RESTART-DCM})$, in which p stands for the rate of the event "non-stuttering" in the treatment group. Non-stuttering was operationalized as ≤1.5% SS [191] on audio recordings made by parents. The audio recordings were scored for %SS by SLPs who were not involved in the study and who were blinded for therapy and measurement moment. Missing audio recordings were replaced by video recordings made in the clinic (n=1 for the LP group; n=2 for the RESTART-DCM group). Further details, along with full details on clinical outcomes, can be found in de Sonneville-Koedoot et al. [213].

The primary outcome for the CUA was HRQOL as measured by proxy versions of the EuroQoL EQ-VAS [113] and the Health Utility Index-3 (HUI3) [111] at baseline and 3, 6, 12 and 18 months post-treatment onset. The EQ-VAS [113] is a visual analogue scale ranging from 0 (worst imaginable health) to 100 (best imaginable health). It was applied for all children under study to derive a direct rating of the child's current health state by parents. The HUI3 [111] is a preference-based measure of HRQOL. That is, health descriptions by parents on a standardized system were linked to empirical values that represent the strength of preferences of the general public for those health states (*utilities*). The HUI3 descriptive system consists of eight domains (vision, hearing, speech, ambulation, dexterity, emotion, cognition and pain and discomfort), with five to six response levels for each domain. The *utilities* range from -.36 (worst imaginable health) to 1 (best imaginable health) [111]. The HUI3 is only applicable for children aged 5 years and older; for younger children no generic preference-based measure of health exists.

HRQOL values were converted into quality adjusted life years (QALYs) by the area under the curve method. This implies that the time spent in a particular health state was multiplied by the value or utility for that health state. If, for instance, a child would be in perfect health (utility of 1) at baseline as well as at 3 months post start treatment, the amount of QALYs generated in the first three months (a quarter) is (1*0.25)=0.25 QALYS. If a child's utility score changes from 0.80 at baseline to 0.90 at 3 months, this would generate: (0.85*0.25)=0.2125 QALYs. QALYs are commonly based on HRQOL utilities (e.g., as obtained by the HUI3). In the current study, HRQOL values as obtained by the EQ-VAS were also applied to estimate QALYs. For a review of the criticism on and advantages of using value-based QALYs (V-QALYs) instead of utility-based QALYs (U-QALYs) in cost-utility analysis we refer to Parkin and Devlin [215].

Resource use and costs

In line with Dutch guidelines for pharmacoeconomic research [216], a societal perspective was adopted. Consequently, direct health care costs, direct non-health care costs and indirect non-health care costs were measured (Table 6.1). Data on the number and duration of treatment sessions, which were expected to have the largest impact on the total costs, were collected directly from registrations by the SLPs. Parental questionnaires included questions on the number of contacts with other healthcare providers to stimulate fluent speech (e.g., physiotherapist, assistance to parents in upbringing), extra material bought by parents to do the treatment at home (e.g., toys, books), transport to the speech clinic (travel distance, volume and means of transport), time invested on homework assignments, hours of absenteeism from paid work due to a clinical visit and duration of a visit to the clinic (including travel time). The unit-cost of a treatment session was derived from a cost analysis conducted by the NZa (Dutch Health care Authority) for the year 2009 [217] and adjusted for 2010 using consumer price indices [218]. In order to account for differences in duration of treatment sessions, the price per minute of therapy was calculated. For the calculation of all other costs, reference prices derived from the Dutch manual for costing [219] were used (Table 6.1). All costs were estimated for the year 2010 and are presented in Euros.

6

TABLE 6.1 Resources and unit prices

Resources	Euro	Source
Direct health care costs		
Treatment for stuttering	34.26 [per half hour]	NZa Cost research
Additional treatment[a]	33 – 36 [treatment session]	Dutch Costing Manual (2010)
Extra material for therapy at home	price given by parents	
Direct non health care costs		
Travel by car	0.20 [per kilometer]	Dutch Costing Manual (2010)
Travel by public transport	ticket price	Dutch Costing Manual (2010)
Homework assignments[b]	12.50 [per hour]	Dutch Costing Manual (2010)
Indirect non health care costs		
Absenteeism from paid work[c]	32.25 [per hour]	Dutch Costing Manual (2010)
Productivity losses for unpaid work[d]	12.50 [per hour]	Dutch Costing Manual (2010)

[a] Additional treatment consisted of speech/language therapy (n=5), physiotherapy (n=2), manual therapy (n=2), osteopathy (n=1) and children's coach (n=1). The price per children's coaching session is valued at the price for a session of 'remedial therapy.'
[b] Time investigated by parents in homework assignments was valued as opportunity costs for housework [219]
[c] To account for differences between age groups and gender in productivity costs, the mean productivity costs for the age group 35 to 40 years were used
[d] Productivity losses for unpaid work consisted of time invested by parents into a visit to the clinic (including travel time). These costs were valued as opportunity costs for housework [219]

Data analysis

Health outcomes, resource use and costs

Intention-to-treat analyses were conducted based on group allocation, regardless of actual intervention received or adherence to the intervention. One child was excluded for all analyses because just after start of treatment it was decided that he had not fulfilled all inclusion criteria. The between-group difference in percentage of non-stuttering children at 18 months, that was applied to calculate the NNT, was compared by a χ^2 test. The level of significance was set at α=.05.

Missing data are common in economic evaluations alongside RCTs and are due to incomplete questionnaires, participants who do not show up at measurement moments or who drop out from the trial. If those with complete data differ from those with incomplete data, bias may influence the results. Analysis of complete cases may also lead to a loss of statistical power. Therefore, missing data in QALYs and costs were multiply imputed. Multiple imputation is a technique is in which each missing value is replaced by at least one simulated value [220,221]. This technique reflects the uncertainty that is inherent when replacing missing data [220,222]. Usually, m=10 is found to be sufficiently large [223]. Thus, we created ten different data sets. We applied the commonly recommended Monte Carlo Markov Chain approach to impute missing data [224]. Variables included in the model were therapy, age at baseline, gender, stuttering severity ratings by clinician and parent, stuttering frequency, scores on the Child Behavior Checklist (CBCL) [194], HRQOL scores and costs. After the multiple imputation procedure, the estimates from the 10 different datasets were pooled using a formula described by Rubin [221].

Because of the skewed distribution of QALYs and costs, bootstrapping (1000 replications) was used to obtain 95% uncertainty intervals around the estimated means of QALYs and costs in both treatment groups. This was done for the 10 datasets that were obtained by multiple imputation. The resulting 10,000 (10x1000) bootstrap replications for the LP and RESTART-DCM treatment were compared and the differences in QALYs and costs were calculated by the 95% uncertainty interval. Resource use was calculated on the observed data before imputation and compared by independent t-tests. The level of significance was set at α=.05.

Cost-effectiveness and cost-utility

To estimate the incremental cost-effectiveness ratio (ICER) of the LP compared to RESTART-DCM treatment, the difference in total costs between treatment groups was multiplied by the number of children that need to be treated by the LP in order to have one additional child defined as non-stuttering at 18 months post-treatment onset. The ICER thus expresses the extra monetary investment needed in order to have one additional child who did not stutter at 18 months.

To estimate the incremental cost-utility ratios (ICURs), results of the bootstrap replications were graphically plotted on cost-utility planes. These planes show the estimated incremental cost per QALY gained and display the uncertainty around the ratios. ICURs were calculated for QALYs based on the EQ-VAS (V-QALYs) and on the HUI3 (U-QALYs). In addition to the cost per QALY taking a societal perspective, we estimated the ICURs including only the direct costs. Acceptability curves (CEACs) were used to present the probability that the LP is more cost-effective than RESTART-DCM treatment for different values of the willingness-to-pay threshold [88]. The WTP threshold represents the maximum cost a decision maker is willing to spend per unit of health outcome (i.e., QALY) gained by a new intervention. Thus, this is the maximum acceptable cost-utility ratio.

Sensitivity analysis

Two sensitivity analyses were performed to test the robustness of the results. In sensitivity analysis one (SA1), costs and QALYs in the second year were discounted. Discounting decreases the value of costs and benefits in after the first year. Costs in the period between 12 and 18 months were discounted with 4% and QALYs achieved in this period with 1.5%. This is in line with Dutch guidelines [216]. In a second sensitivity analysis (SA2), the impact of imputation of missing data was tested by solely including children that completed the trial (i.e., complete case analysis).

Statistical analysis

Analyses were performed using SPSS version 20 (Armonk, NY: IBM Corp.). CEACs were constructed using MS Excel 2010.

RESULTS

Participant flow and baseline characteristics

The participant flow of the 199 children randomized to the LP (N=99) and RESTART-DCM treatment (N=100) is presented in Figure 6.1. In total, 11 children missed one or more measurement moments and 21 children (11%) dropped out from the study. Reasons for not completing the trial included relocation (n=4), families being unavailable (n=6), lack of motivation for participation because of fluent speech (n=2), family problems (n=6), and one SLP who stopped participating in the trial shortly after inclusion of children (n=3). In total, 177 children completed the 18-month assessment (86 in the LP group and 91 in the RESTART-DCM treatment group). Children who completed the study and children who dropped out did not significantly differ on baseline characteristics. For 191 children, at least one outcome measurement after start of treatment and at least one cost booklet were available. They were therefore included in the economic evaluation. Baseline characteristics of all randomized children are presented in Table 6.2.
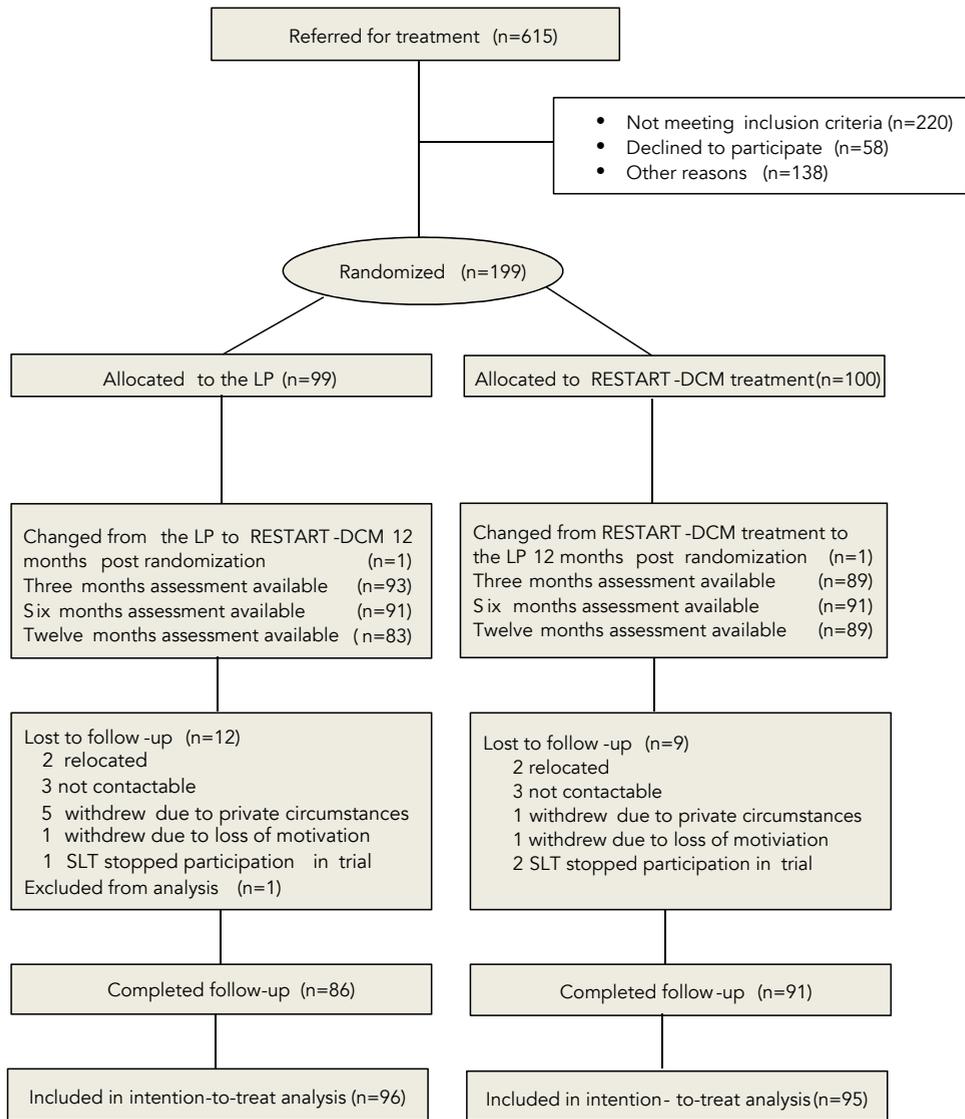
FIGURE 6.1. RESTART trial enrollment, randomization, and assessment

TABLE 6.2 Baseline characteristics of participants by treatment group

Characteristic	Lidcombe Program (n=98)[a]	RESTART-DCM treatment (n=100)[a]
Age in months, median; mean (SD)	51.0; 51.5 (9.5)	52.0; 54.1 (11.1)
Age in years		
3[b]	41 (41.8)	37 (37.0)
4	39 (39.8)	31 (31.0)
5-6	18 (18.4)	32 (32.0)
Male	68 (69.4)	70 (70.0)
SSI-3 score		
mild[c]	32 (32.7)	31 (31.0)
moderate	47 (48.0)	51 (51.0)
severe	19 (19.4)	18 (18.0)
% SS, median; mean (SD)[d]	4.9; 6.2 (4.4)	4.0; 5.3 (4.3)
Time since onset		
6-12 months	43 (43.9)	45 (45.0)
13-18 months	25 (25.5)	22 (22.0)
19+ months	30 (30.6)	33 (33.0)
Family history of persistency[e]	45 (45.9)	45 (45.0)
Family history of recovery[e]	27 (27.6)	25 (25.0)
Prior treatment for stuttering	8 (8.2)	6 (6.0)

[a] Data are shown as No. (%) unless specified otherwise
[b] One child in the LP group was 2.11 years at time of inclusion
[d] Children with a stuttering frequency < 3% in the therapy setting but ≥ 3% in the home setting were included in the group "mild stuttering"
[d] For one child in the RESTART-DCM group %SS on baseline was not available
[e] For one child in the LP group information on family history of stuttering was not available

Health outcomes

The percentage of children who did not stutter at 18 months has been reported in de Sonneville-Koedoot et al. [213] and were respectively 76.5% ((65/85)*100) in the LP group versus 71.4% ((65/91)*100) in the RESTART-DCM group. The difference of 5.1% was statistically non-significant ($\chi^2(1)=0.579$, $p=.45$). The NNT was (1/0.051=) 20; i.e., if 20 children are treated by the LP instead of RESTART-DCM treatment, on average one more child will be defined as non-stuttering after 18 months.

Table 6.3 lists the EQ-VAS and HUI3 scores at baseline and at follow-up. Both groups started with relatively high mean HRQOL-scores at baseline and showed a slight increase in scores over time. The mean EQ-VAS score in the LP group was lower at baseline and higher at the end of the 18-month period than in the RESTART-DCM group; V-QALYs gained in this group were therefore slightly and significantly higher (LP: 1.36; RESTART-DCM: 1.34; difference: 0.018 (95% CI: 0.008 to 0.027)). The effect size was small (Cohen's d=0.17;

[117]. For children aged 5 years and older, U-QALYs based on the HUI3 in 18 months were 1.38 for the LP group and 1.37 for the RESTART-DCM group. The difference in U-QALYs did not reach statistical significance (0.013; 95% CI: -0.006 to 0.03).

TABLE 6.3 HRQOL over the 18-month follow-up period

	EQ-VAS		HUI3	
	LP (n=98)	RESTART-DCM treatment (n=100)	LP (n=18)	RESTART-DCM treatment (n=33)
	mean (SD)	mean (SD)	mean (SD)	mean (SD)
Baseline	88.01 (9.86)	88.34 (9.69)	0.88 (0.13)	0.88 (0.11)
3 months	89.57 (10.24)	88.60 (10.61)	0.92 (0.12)	0.88 (0.15)
6 months	91.05 (7.40)	89.63 (8.85)	0.89 (0.13)	0.90 (0.11)
12 months	91.27 (9.15)	89.64 (11.96)	0.95 (0.09)	0.94 (0.10)
18 months	91.42 (8.81)	90.56 (9.60)	0.95 (0.10)	0.94 (0.10)

Abbreviations: EQ-VAS: EuroQol Visual Analog Scale; HUI3: Health Utility Index-3

TABLE 6.4 Mean resource use per child for both treatment groups during the 18-month follow-up period and the percentage of participants using this resource

Resources	LP (n=98)			RESTART-DCM treatment (n=100)					
	n	mean	SD	n	mean	SD	mean	95%	UI
Direct health care costs									
Stuttering treatment [sessions]	98	22.2	11.2	97	19.5	10.3	2.7	-.3	5,8	.08
Stuttering treatment [minutes]	95	1198	656	93	1080	584	117.7	-61.1	296,6	.20
Additional treatment [sessions][a]	95	.38	1.60	92	.23	1.42	.1	-.3	,6	.52
Direct non health care costs									
Travel by car [kilometers]	95	453.9	543.5	94	305.9	370.8	148.0	14.5	281,5	.03
Homework assignments [hours][b]	92	59.9	62.3	92	62.6	58.3	-2.7	-20.3	14,9	.76
Indirect costs									
Absenteeism from paid work [hours]	95	1.39	4.27	93	1.75	5.24	-.4	-1.7	1,0	.61
Productivity losses for unpaid work [hours][d]	95	32.5	22.8	93	25.5	19.6	7.0	.8	13,1	.03

Results in this table are based on analysis of observed data, before imputation.
[a] Additional treatment as specified in Table 6.1
[b] For four children the number of hours invested in homework estimated by parents was unrealistically high, therefore these were defined as outliers and excluded from analyses
[c] Productivity losses for unpaid work consisted of time invested by parents into a visit to the clinic (including travel time)

Resource   use

Table 6.4 shows the mean utilization of resources in both groups over 18 months. Children in the LP group had slightly more treatment sessions (22.2) than children in the RESTART-DCM group (19.5). The difference was not statistically significant (p=.08). The larger travel distance and higher productivity losses for unpaid work (i.e., time invested by parents into a visit to the clinic, including travel time) in the LP group compared to the RESTART-DCM group just reached statistical significance (p=.03 for both comparisons).

Costs

Detailed cost comparisons are listed in Table 6.5. The average costs refer to the costs per child. The mean total costs for the LP group were significantly higher than for the RESTART-DCM group (difference €168, 95% CI: €61 to €277), but the effect size was small (Cohen's d=0.14). The higher costs were largely due to higher treatment costs and thus to a greater number of treatment sessions for the LP group.

TABLE 6.5 Mean costs in Euros per child for both groups and difference between groups during the 18-month follow-up period

	LP (n=98)				RESTART-DCM treatment (n=100)				mean difference		
	mean	s.e.	95%	UI	mean	s.e.	95%	UI	mean	95%	UI
Direct health care costs	1474	719	1430	1519	1334	645	1294	1374	140	80	201
Stuttering treatment costs	1422	693	1379	1465	1287	618	1249	1326	135	76	192
Additional treatment costs[a]	26	85	21	32	27	104	21	34	-1	-10	7
Extra material costs for therapy at home	25	41	23	28	19	29	18	21	6	3	9
Direct non health care costs	1177	728	1132	1223	1215	779	1167	1262	-38	-103	29
Travel costs (includes travel by car and by public transport)	111	109	105	118	79	71	75	84	32	24	40
Costs associated with time invested in homework assignments	1065	716	1022	1111	1136	760	1089	1183	-70	-134	-6
Total direct costs	2650	1047	2587	2716	2549	1130	2480	2619	102	6	197
Indirect costs	548	259	532	565	483	223	470	497	65	44	87
Costs associated with absenteeism from paid work	58	137	50	67	75	184	63	86	-17	-31	-3
Costs associated with productivity losses for unpaid work	490	248	475	507	409	206	396	421	82	62	102
Total costs	3199	1203	3125	3275	3032	1272	2955	3109	168	61	277

[a] Additional treatment as specified in Table 6.1

Cost-effectiveness

Table 6.6 presents the difference in costs and effects as well as the incremental cost-effectiveness ratio (ICER) and cost-utility ratios (ICURs) from a societal perspective and if only direct costs are included. Based on the NNT of 20 and the difference in costs of €168, the estimate of the incremental cost per NNT was (20 * €168 =) €3360. Thus, an extra investment of €3360 would result in one more child classified as non-stuttering with the LP compared to RESTART-DCM treatment. The incremental cost per V-QALY estimate (based on the EQ-VAS) was €10413 for the LP compared with RESTART-DCM treatment. For children aged 5 to 6 years, the ICUR based on the HUI3 amounted to €18617. Uncertainty around these ratios is illustrated in the cost-effectiveness planes (Figure 6.2 and 6.3). Each of the 10000 points represents a pair of incremental cost and effect value replicates for the LP group compared with the RESTART-DCM group, produced by the bootstrapping technique. Most cost-effect pairs for the ICUR are located in the northeast quadrant (99.8% of the ratios for cost per V-QALY and 89.2% for cost per U-QALY), indicating that the LP is associated with a higher HRQOL after 18 months and higher costs compared to RESTART-DCM treatment. If only direct costs were included, the ICER was €2040, the ICUR based on V-QALYs €6373 and the ICUR based on U-QALYs €12110 (Table 6.6). Again, most cost-effect pairs were located in the northeast quadrant: 98.2% and 75.1%, respectively (Figures A6.1 and A6.2 in the Appendix).

TABLE 6.6 Difference in mean health outcomes and costs (with 95% confidence intervals) and incremental cost-effectiveness ratios (ICERs) and cost-utility ratios (ICURs) of the LP group compared to the RESTART-DCM treatment group

	Societal perspective	Including only direct costs
Health outcomes		
Percentage of children who do not stutter at 18 months	5.1 (-0.08 to 0.18)	
V-QALY (EQ-VAS)	0.018 (0.008 to 0.027)	
U-QALY (HUI3)	0.013 (-0.006 to 0.03)	
Costs (€)		
For all children	168 (61 to 277)	102 (6 to 197)
For children aged 5 to 6 years	228 (2 to 454)	90 (-102 to 279)
Cost-effectiveness ratios		
ICER: Costs per NNT	3360	2040
ICUR: Costs per extra V-QALY gained	10413	6373
ICUR: Costs per extra U-QALY gained	18617	12110

Abbreviations: EQ-VAS: EuroQol Visual Analog Scale; HUI3: Health Utility Index-3
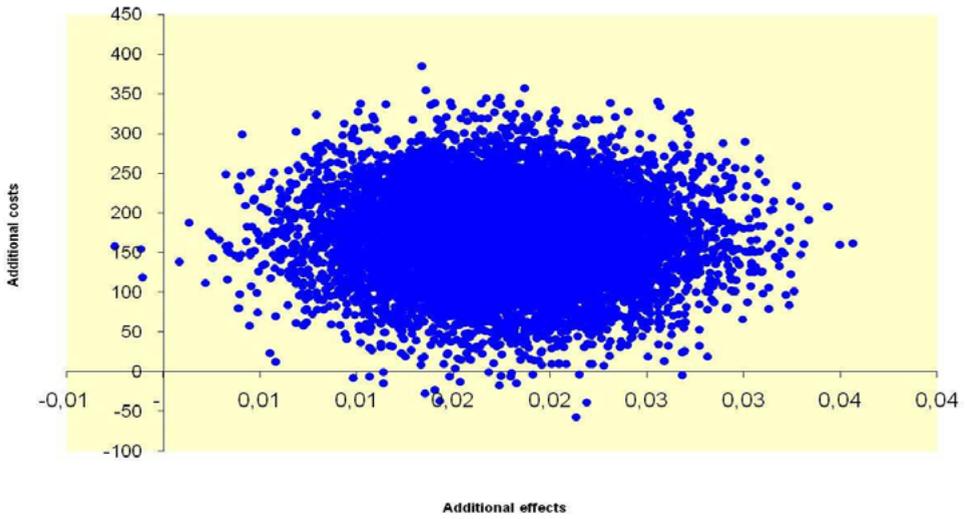
FIGURE 6.2 Cost-effectiveness plane for the difference in V-QALYs (based on EQ-VAS) gained in 18 months
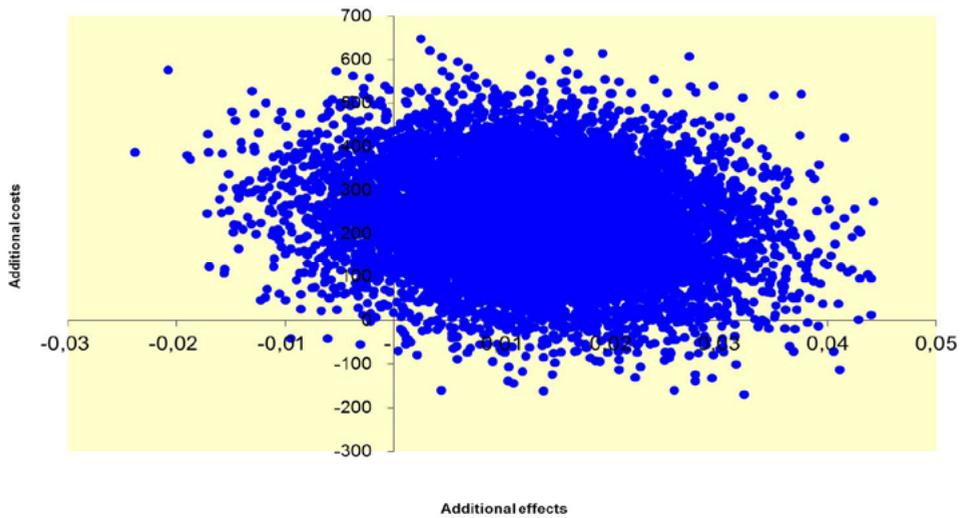
FIGURE 6.3 Cost-effectiveness plane for the difference in U-QALYs (based on HUI3) gained in 18 months

Figure 6.4 presents the CEACs for the ICUR based on V-QALYs and on U-QALYs. It shows the probability that the LP is cost-effective compared with RESTART-DCM treatment, at different levels of WTP for a QALY. At a societal WTP of €20.000 per V-QALY, the probability is 0.95. At a WTP of €35.000 per V-QALY, the probability rises to 0.99. For the WTP for a U-QALY these probabilities are 0.55 and 0.74, respectively.
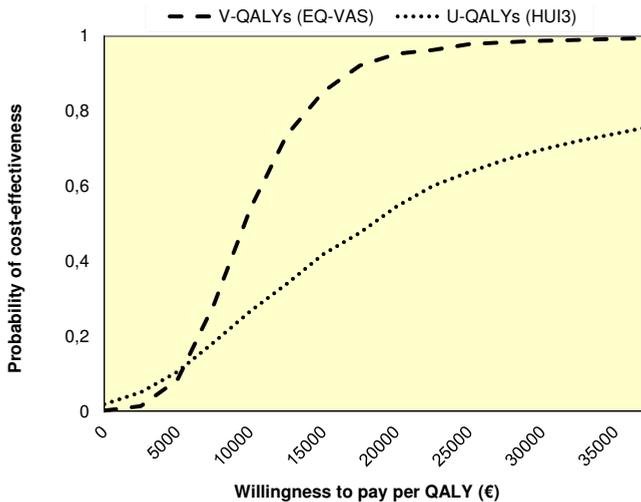
FIGURE 6.4 Acceptability curves for the LP compared to RESTART-DCM treatment

Sensitivity analysis

Results from SA1, in which costs and effects were discounted, were comparable with the results from the main analysis (data not shown). Complete case analysis (SA2) was done for the outcomes percentage of children classified as non-stuttering (LP: n=85, RESTART-DCM treatment: n=91) and EQ-VAS (LP: n=61, RESTART-DCM treatment: n=67). For the outcome HUI3, complete data was only available for 21 children in the LP group and 12 children in the RESTART-DCM group. Therefore, complete case analysis for this outcome was not performed. For the outcome percentage of non-stuttering children, the mean difference in effect for complete case analysis was similar compared with the main analysis (5.0% in favor of the LP, 95% CI: 1.1 to 8.9). The costs for the LP group were €2591 versus €2852 for the RESTART-DCM group. The mean cost difference of €260 (95% CI: €138 to €383) was larger than found in the main analysis. This resulted in a higher cost per NNT estimate than established in the main analysis: (20 * €260 =) €5200. For the health outcome V-QALYs, the mean difference in effect was 0.01 (95% CI: 0.002 to 0.02). Mean total costs in the LP were €2975 and in the RESTART-DCM group €2927 (difference: €48; 95%CI: -€67 to €169). This resulted in a lower estimate of cost per V-QALY than in the main analysis: €5410 (i.e., extra costs of €5410 for each V-QALY gained).

DISCUSSION

This is the first study to report both costs and effects of stuttering therapy in preschool children. It was demonstrated that, over the first 18 months after treatment onset, there is a high probability that the LP is slightly more costly than RESTART-DCM treatment but also leads to slightly better health outcomes. Differences in total costs and V-QALYs (quality-adjusted life years based on the EQ-VAS) were statistical significant but effect sizes were small; percentage of children who did not stutter at 18 months and U-QALYs (based on the HUI3, only applicable for children aged 5-6 years) were not statistically significant between groups. Due to the marginal between-group difference in cost (€168), small differences in effects resulted in relatively low cost-effectiveness and cost-utility ratios. The costs per extra child classified as non-stuttering at 18 months with the LP compared to RESTART-DCM treatment were €3360. The costs to obtain extra improvement in HRQOL by the LP compared to RESTART-DCM treatment were €10413 per extra V-QALY and €18617 per extra U-QALY.

Study design
The RESTART-trial employed a practical study design, ensuring a high external validity and allowing for economic evaluation [220]. Unfortunately, the study did not provide an opportunity to evaluate the cost-effectiveness of stuttering treatment after the time horizon of 18 months. Treatment was not completed in 28% of children allocated to the LP and 35% of children allocated to RESTART-DCM treatment [213], so total treatment costs are expected to incline after 18 months. Furthermore, follow-up data is needed to confirm the difference in percentage of non-stuttering children on longer term and to establish long-term HRQOL effects of stuttering. Our study into the HRQOL of adults who stutter has shown that mainly severe stuttering in adulthood is associated with a lower HRQOL [91]. Thus, in order to accurately predict the long-term comparative cost-effectiveness of stuttering treatment, data on stuttering severity, follow-up costs and health effects need to be incorporated, for instance in a statistical model. Anticipating on this, the slightly higher percentage of children on treatment at 18 months in the RESTART-DCM group suggests that the difference in treatment costs might further reduce. Moreover, if treatment effects would remain similar over life time, the 18-month time horizon possibly underestimates the difference in QALYs gained between therapies. For both reasons we consider our current estimate of the cost-effectiveness of the LP compared with RESTART-DCM treatment conservative; more favorable results for the LP would most likely have been obtained by applying a longer time horizon.

6

Health-related Quality of Life

Mean HRQOL values were high at baseline and increased only slightly during follow-up, resulting in almost comparable between-group EQ-VAS and HUI3 scores at 18 months. While this suggest that stuttering at preschool age does not generally impair HRQOL, several limitations to our assessment of HRQOL must be noted. First, in economic evaluations from a societal perspective, preference-based health state values are preferred over ratings of health states by patients or proxies (i.e., parents). Currently, however, no HRQOL instrument is available for QALY computation for children under the age of five years. This made the use of the proxy EQ-VAS version indispensable. It also yielded the only generic outcome applicable for all children under study. We transformed EQ-VAS scores into values applicable for QALY computation, which has also been done in other studies (e.g., [225]). Second, the overall increment in mean EQ-VAS score over 18 months in both groups was slightly lower than the increment in mean HUI3 score. This result may be attributed to limited responsiveness of the EQ-VAS to changes in mild problems, which arise as a result of the end-of-scale bias (see [88], and [91] for a discussion on this bias in relation to stuttering). Moreover, it is not evident that stuttering is perceived as a health problem by parents [226]. However, since the HUI3 scores at 18 months showed almost no differences in HRQOL between treatment groups for children aged 5 and 6 years and the between-group differences in V-QALYs were comparable to the difference in U-QALYs, we believe that the use of the EQ-VAS (and, accordingly, the computation of V-QALYs) was warranted. A last limitation related to the HRQOL measurement is the use of proxies in general. It is known that the adequacy of proxy ratings may be confounded by characteristics such as parental emotions and stress [121-123].

Costs

Examination of the costs revealed that the total difference of €168 was mainly due to direct treatment costs being higher for the LP group (€1422 versus €1287). This was anticipated, since the LP is known to require a relatively long maintenance phase. It appeared that, although a slightly higher percentage of children in the RESTART-DCM group was on treatment at 18 months, children who terminated treatment in the LP group had on average followed more treatment sessions than children in the RESTART-DCM group. Few parents reported use of other healthcare services and costs for practice material were low for both groups. Time spent on homework assignments was less in the LP group than in the RESTART-DCM group (59.9 versus 62.6 hours). This probably reflects a difference in therapy method. RESTART-DCM treatment requires parents to spend 15 minutes a day on homework, while the LP gradually integrates the use of contingencies in daily activities. All parents may have had difficulties in estimating time spend on homework assignments, as indicated by the relatively high percentage of missing data for this resource use.

Exclusion of the homework related costs did not lead to different results (data not shown). Travel costs per session were higher for the LP group. Although this result was obtained coincidentally, it might also be the case in real world, since fewer Dutch SLPs are trained in the LP than in RESTART-DCM treatment. Thus, parents probably need to travel a longer distance in order to receive treatment according to the LP. The larger number of treatment sessions and travel costs associated with the LP resulted in higher total costs associated with productivity losses for this group, and thus in higher indirect costs. Exclusion of the indirect costs did not alter the conclusion that the LP was more costly.

Cost-effectiveness

The benefit for the LP over RESTART-DCM treatment in terms of percentage children classified as non-stuttering and amount of QALYs gained was minimal, but so were the extra costs. Whether or not the LP can be regarded a cost-effective approach depends on the maximum willingness-to-pay for such small improvements in health. Extra costs to gain one extra case of non-stuttering were estimated at €3360. There is, however, no consensus about the maximum WTP for a clinical outcome as "non-stuttering". Therefore, health economists prefer to express treatment benefits in a generic outcome such as the QALY. When evaluated against the threshold of €20.000 per QALY that is often suggested as an upper limit in the Netherlands, the results on the U-QALYs indicate a probability of 55% that the LP is a cost-effective therapy compared with RESTART-DCM treatment. This amounted to 74% using a threshold of €35.000. Based on V-QALYs, the LP would be considered cost-effective compared to RESTART-DCM at lower levels of WTP: 95% certainty for a WTP threshold of €20.000 and 99% for a threshold of €35.000 per QALY. Complete case-analysis resulted in even a lower ICUR, but this ratio is probably biased as a result of the small sample size. Different parents, different SLPs, and different health care funders will have individual perceptions and judgements on the clinical and financial value of these results. In addition, health care resources used and unit costs for other countries can differ somewhat, due to variation in organization and financing of health care. For Dutch daily practice, the results clearly indicate that the LP can be considered a cost-effective alternative to RESTART-DCM treatment.

An interesting future topic of research, that could result in potential cost savings, is to reconsider the relation between treatment duration and health outcomes. In both groups, quite a large number of children were still on treatment at the end of the trial, including about 20% of children who were classified as non-stuttering. This might suggest that criteria for terminating treatment in daily practice are too stringent. It can be questioned if active monitoring may not be as sufficient as active treatment in order to accomplish the goal of barely noticeable stuttering that both treatments strive for, a topic that is however beyond the scope of this study.

6

CONCLUSIONS

In conclusion, differences in effects and costs between the LP and RESTART-DCM treatment were small and cost-effectiveness and cost-utility ratios were in favor of the LP. This indicates that the LP is a good alternative to RESTART-DCM treatment in Dutch primary care.

ACKNOWLEDGMENTS

APPENDICES CHAPTER 6

FIGURE A6.1 Cost-effectiveness plane for the difference in V-QALYs (based on EQ-VAS) gained in 18 months, including only direct costs
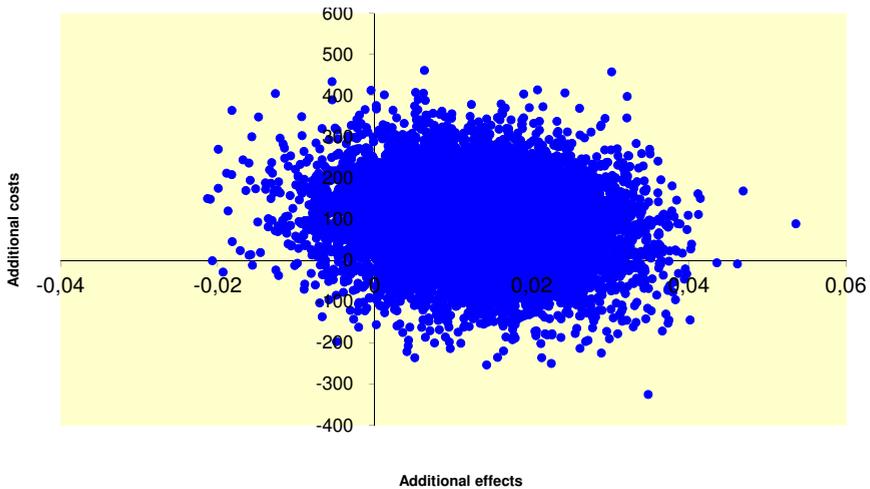

FIGURE A6.2 Cost-effectiveness plane for the difference in U-QALYs (based on HUI3) gained in 18 months, including only direct costs

7

Perspectives of clinicians involved in the RESTART-study: outcomes of a focus group

With Samantha Adams
Elly Stolk
Marie-Christine Franken

ABSTRACT

Purpose To explore the attitudes and beliefs of speech-language pathologists (SLPs) with regard to the Lidcombe Program (LP) and Demands and Capacities (DCM) based treatment, and how these might have changed as a result of participating in the RESTART-study.

Methods A focus group meeting with 13 SLPs was organized. The discussion was structured using questions on therapy preference, attitudes about and explicit comparison of both treatments and treatment manuals, and learnings of trial participation.

Results Four main themes were identified. Firstly, a change in attitude toward treatment choice was observed. Secondly, this was related to a change in beliefs about the potential of both treatments. Thirdly, aspects of the treatments regarded as success factors were considered. Lastly, learning outcomes and increased professionalism as a result of participating in the RESTART-trial were discussed.

Conclusions This study showed how attitudes and beliefs of SLPs with regard to the LP and DCM based treatment evolved during a randomized trial. This work increases our understanding of the role of attitudes and beliefs in the uptake and utilization of therapies and demonstrates the importance of collecting qualitative data. Results and recommendations should prove of value in implementing the RESTART-trial results and in training of SLPs.

INTRODUCTION

Early intervention is currently considered best practice in therapy for children who stutter [10,184,227]; however, there is limited evidence as to the relative effectiveness of therapies for preschool children who stutter. A recent multicenter Randomized Controlled Trial (RCT) conducted in the Netherlands compared the effectiveness and cost-effectiveness of two widely applied therapies: treatment based on the Demands and Capacities Model (DCM) and the Lidcombe Program for early stuttering intervention (LP). The RESTART-study (acronym for the *R*otterdam *E*valuation study of *S*tuttering *T*herapy- *A R*andomized *T*rial) was conducted in the period September 2007 to December 2011. The study included a total of 199 preschool children who stutter and 24 speech-language pathologists (SLPs). Although a RCT is generally perceived to be the most highly regarded research method, it is also acknowledged that not all kinds of questions can be addressed by this method [228]. Moreover, daily clinical practice is known to be greatly influenced by subjective factors, such as a clinician's own perceptions and experiences [229].

With DCM based treatment being the standard treatment in a large part of the world, and an increasing number of clinicians becoming experienced in the LP, a growing number of SLPs are nowadays in the position to make a choice between treatments. However, little research data is available as regards clinicians' views on both treatments, factors governing their therapy selection process, or their treatment delivery. This is all the more relevant since these treatments are premised on different ideologies and, accordingly, apply a fundamentally different therapy approach [45,46]. DCM based treatment uses an *indirect* approach, based on the theoretical model that stuttering is a multifactorial disorder [52-54] with physical, linguistic, psychological, and/or environmental factors influencing the onset and development of stuttering. Accordingly, treatment focuses on manipulating child related and environmental factors assumed to influence the child's speech fluency. Contrary to DCM based treatment, the LP directly targets the child's speech fluency. Treatment is not explicitly based on a theory of the onset and development of stuttering, but assumes that manipulating the child's speech production by using operant conditioning procedures will lead to increased fluency. One of the essential components of the LP is the use of parent verbal contingencies for stutter-free speech and for stuttering [48]. Thus, basically SLPs can choose between a more holistic and a more technical-behavioral approach to treat preschool children who stutter. Presumably it is their experiences, views and opinions that will shape their decisions in this matter.

Perceptions of clinicians or other stakeholders on stuttering treatment for children have only marginally been addressed in literature. There are no studies available on the perspectives of clinicians for DCM based treatment, while only three explorative studies addressed the

7

experiences of Australian SLPs with the LP. A study by Harrison, Ttofari, Rousseau, and Andrews [230] showed that the main reasons for SLPs consulting an expertise center were limited progress with the LP, or difficulties with the treatment when working with relatively unusual clients. Studies by Rousseau, Packman, Onslow, Dredge, and Harrison [231] and, more recently, O'Brian et al. [232], revealed that around half of the clinicians surveyed did not deliver the LP in the prescribed format. These examples illustrate that engagement of clinicians in the evaluation of a therapy can provide improved understanding of the uptake and utilization of that treatment. Evaluation of SLPs perceptions is also critical with regard to the translation of research evidence into practice. Strong personal preferences, non-adherence to manuals or perceived barriers to treatment implementation can all influence how study results will find their way into daily practice. Since therapists participating in a trial could be regarded as "early adopters" [233], it is of particular relevance to obtain insight into their attitudes, beliefs and experiences.

Therefore, it was decided to explore of the attitudes, beliefs and experiences of the clinicians participating in the RESTART-trial by means of a focus group study. By explicitly using group interaction as methodology, a focus group interview is an ideal way to generate qualitative good data about participants' experiences and opinions on a defined set of issues [234,235]. The interaction between participants enables an in-depth discussion of issues, which can therefore lead to the generation of rich data [236]. The purpose of this focus group study was to explore the attitudes and beliefs of SLPs participating in the RESTART-trial with regard to the LP and DCM based treatment (RESTART-DCM based treatment[1]), and how these might have changed during their experiences of participating in the trial. The questions used for the focus group discussion are outlined in the methods section.

METHODS

Participants

Twenty-four SLPs throughout the Netherlands participated in the RESTART-study. All had received training for DCM based treatment in their regular education and in addition had followed the formal LP training. In the final phase of the project (once the inclusion of children had ended and the final cohort was completing the last measurements), the SLPs still involved in the project (n=20) were invited to attend the focus group meeting. Thirteen SLPs participated in the focus group. All but one was engaged in the RESTART-study from the beginning of the trial. Demographic characteristics are displayed in Table 7.1.

---

1   Whenever we refer to treatment based on the Demands and Capacities Model as it has been adapted for use in the Netherlands, we describe it as 'RESTART-DCM treatment' [188].

TABLE 7.1 Demographic characteristics of participants

Characteristic	Mean (range)
Female; No. (%)	13 (100%)
Age in years	50 (43-63)
Years of experience DCM based treatment	14.5 (7-21)
Years of experience LP	3.0 (1.5-7.0)
Work in private practice; No. (%)	13 (100%)

Note. Data are shown as mean (range) unless specified otherwise

Procedure

Participation in the focus group was voluntary. To compensate for loss of income and travelling expenses, participants received a small fee. The focus group was held at the location as where participants had had regular quarterly meetings during the trial and lasted approximately 2 hours and 45 minutes (with a 15 minute break). A protocol outlining the relevant themes and order in which they were to be discussed was developed by the authors. Since participants had met each other in quarterly discussion meetings during the previous 3 years and 10 months of the trial, the researchers and participating SLPs were quite familiar with each other. While this created a comfortable setting for discussion, it also brought the risk of bias in the data (for example, participants would not feel comfortable expressing their true feelings about the project, or presumed shared understanding might inhibit expressing shared experiences). Therefore, the focus group was moderated by an experienced senior researcher, unknown to the participants. She had been informed about the therapies and RESTART-study by the other authors. Assistant interviewers were the first and last author. The first author is the primary investigator; the last author is the initiator of the RESTART-study and also a SLP.

In the focus group meeting, the interviewers adopted the role of process facilitators (checking time, ensuring the protocol was followed, asking follow-up questions, etc.) and were conscious of being open to accommodate both positive and negative experiences expressed by the informants. They were also conscious of the importance of encouraging the participants to express views of a different nature, including both positive and negative views. Participants who were relatively quiet were encouraged to contribute to the discussion by the interviewers particularly inviting them to respond to the question at hand. After discussion of each question, and before moving on to the next question, the moderator explicitly asked if anyone wanted to add anything else to the discussion. Where appropriate after discussion of a specific question, the moderator counted and recorded the number of participants holding a particular view.

The set of questions a-priori formulated for the discussion covered the following main themes:

- Therapy preference -LP versus RESTART-DCM- *prior to* the RESTART-study. What determined their choice and in which situations did the therapists choose to deviate from the therapy they generally preferred?
- Therapy preference *in the final stage* of the RESTART-study.
- Overall impression of and attitudes to both therapies *prior* to the RESTART-study, and how these may have changed *during the trial* (e.g., flexibility, level of difficulty, goals that could be achieved).
- Attitudes to both treatment manuals (e.g., feasibility, clarity, completeness).
- Explicit comparison of treatments and treatment manuals (e.g., components that are unique and/or essential for therapy success, common elements, time investment for SLPs).
- Experiences with the trial and recommendations for follow-up research.

Data analysis

The interview was recorded (audio only) with permission of the respondents and transcribed verbatim by the first author. The first and second author independently coded the transcripts (manually) in two phases following Creswell's [237] ten steps for qualitative data analysis. The first coding phase was an inductive phase in which the transcript of the focus group was reviewed for recurring themes related to the questions addressed in the focus group discussion. Each author made a coding list of primary and secondary themes that emerged from the data. These were discussed to standardize terminology and ensure that both coders applied the terms in a consistent manner. The transcript was then revisited and coded with the standard terminology and then discussed by the authors a second time. The predominant themes that emerged were then written in a thematic content analysis by the first and second author. To ensure investigator triangulation, this analysis was checked and commented on by the third and fourth authors (one of whom had been present at the focus group meeting). This served as internal peer review of the interpretation of the data [237]. The focus group quotes were translated from Dutch to English by the second author (a native speaker of English, with certified fluency in Dutch) and checked for accurate use of nuance and diction by the first and last authors (native Dutch speakers). All data (recordings and transcripts) were treated confidentially in accordance with institutional policies for storing and analyzing data. All personal identifiers that could link statements to respondents have been removed in this manuscript; participants are identified by a number.

RESULTS

Four main themes emerged from the analysis of the focus group interview. The themes are described in Table 7.2 and summarized and exemplified with quotes below.

TABLE 7.2 Overview of findings

Themes	Subthemes
(1) Changing attitudes towards the choice of treatment	• Strength of specific preferences had decreased • More nuance and flexibility • Parent's preferences and child-specific criteria play greater role
(2) Changing beliefs about the potential of the treatments	• Increased awareness of structure of RESTART-DCM treatment • More flexibility in applying RESTART-DCM treatment • Slightly more positive about the LP than before • Appreciation of clear structure of the LP • Appreciation of rapid effect of the LP • Limited possibilities of LP in case of lack of progress

(3) Attitudes towards success factors of the treatments	• Key-factors RESTART-DCM: slowing down tempo of speech and communication, training in turn taking, changing aspects of parent-child communication (e.g., applying basic rules for communication
	
	• Key-factors LP: Verbal contingencies, structure of stage 2 for monitoring fluency maintenance
	
	• Similarities between RESTART-DCM treatment and the LP:
	
	  1. Initial one hour weekly sessions;
	
	  2. Parents are trained to practice with their child at home in one-to-one situations;
	
	  3. 15 minutes of practice every day;
	
	  4. Overall intensity of time investment for parents is the same;
	
	  5. Importance of pleasant parent-child interaction;
	
	  6. Aim is to increase fluent speech;
	
	  7. SLP needs to constantly refine the treatment to enhance fluency.

(4) Learning outcomes and increased professionalism as a result of participating in the RESTART-trial	• Now start clinical contact with a comprehensive assessment of the child's speech and language • Increased expertise with RESTART-DCM treatment due to working with the manual • Regular meetings with clinicians led to increased learning

7

1. Changing attitudes toward the choice of treatment

Seven therapists reported they had a preference for RESTART-DCM treatment before the start of the trial, five had a preference for the LP and one said she had no preference at all. Although four therapists reported a slight change of general preference from one therapy to the other, all therapists revealed that, in contrast to the beginning of the trial, their preferences and decisions to use one method or the other now reflected more nuance and flexibility; the strength of a specific preference had decreased.

For the majority of the therapists, their former preferences were mainly constituted by personal perceptions, while post-trial, other factors had become more important in selecting a treatment. For example, participants who preferred RESTART-DCM treatment at the start of the trial, mentioned specific perceptions of this treatment, for example it was familiar to them and more comprehensive than the LP. Participants who preferred the LP, on the other hand, most often mentioned the clarity and structure of the program.

> I (…) It just 'suited' me and I was familiar with it. I could use it in any situation. (resp 12, explaining her former preference for RESTART-DCM treatment)

At the final stage of the trial, therapists indicated being open to starting with either therapy and giving more weight to what suits a specific child and family.

> I actually no longer really have a preference. I think that I give the parents and the child more room to decide and that the choice depends less on me. (resp 1, explaining the change experienced during the trial)

Thus, when selecting a treatment program for a preschool child who stutters, in general the therapists at the end of the trial applied different criteria to their starting criteria. All participants indicated that they discuss both therapies with parents and let preferences of parents play an important role in the choice for a specific therapy. Moreover, several therapists mentioned greater use of child-specific criteria on which they now base their choice. Six therapists mentioned that they were inclined to start with RESTART-DCM treatment in younger children because of the more indirect approach of the therapy. Other child-specific factors mentioned that influence choice of therapy were stuttering severity (two participants said that in case of mild stuttering they would start with RESTART-DCM treatment), abilities and temperament of the child (two participants said that they who would start with the LP in precocious children), speech and language problems (as an indication to start with RESTART-DCM based treatment; mentioned by one participant) and the child's attitude regarding his/her speech abilities (mentioned three times). However, therapists did not agree on *how* the child's attitude influences their current therapy choice:

in case of a negative speech-attitude, two SLPs would prefer RESTART-DCM treatment while one SLP would prefer the LP.

2. Changing beliefs about the potential of the treatments

The focus group discussion brought up that, for the majority of therapists, changes in attitudes toward treatment selection as described above were related to subtle changes in beliefs about the potential of both therapies. Participants felt that participating in the trial broke through preconceptions about the therapies. That is, the aspect of randomization forced them to abandon their initial attitudes about what would be the best treatment for a specific child, and the results they achieved with either therapy showed them more clearly the potential of each therapy in its own right. We discuss each of these in turn.

With regard to RESTART-DCM treatment, the most important theme that emerged during the discussion was that the therapists had an increased awareness of how the therapy was structured. This was a result of working with the treatment manual that was constructed for the trial, as well as of shared insights and experiences from other therapists during the quarterly meetings with the group of participating therapists. Some therapists specifically mentioned the clear structure of the theoretical framework of the RESTART-DCM treatment, while others pointed to the practical tools it affords and the basis it provides for explaining the therapy to parents.

> In the past, I never really felt that I had a theoretical basis for what I was doing. It was piecemeal – I pulled one part from one book and one from another and shaped these into some sort of coherent approach. Now, I have a integral approach in my head and can select what I need. (resp 2)

> I have the feeling that I am now more convincing as a therapist in my conversations with parents. I have something to offer. (resp 12)

However, the structure of the RESTART-DCM treatment manual did not lead to rigidity in adopting the DCM approach. Rather, therapists mentioned an increase in flexibility due to the structure and clearness of the manual.

> DCM based treatment is much clearer now, partly because of the protocol and the discussions we had. Where it used to be fairly open with regard to what you can do with DCM – you could call almost anything DCM – or at least, that seemed to be the general feeling that I experienced, now I have a clearer perception of what it is. So, ok, that is how we approach it, those are the things that we are looking at. (resp 5)

As this participant mentioned, the RESTART-DCM treatment manual was instrumental in providing structure to the treatment. It was described as clear and full of practical tips, but the participants also recognized that not all aspects of therapy can be covered in a manual. It was a useful starting point, but also provided therapists the room to shape their own approach in practice. How they did this, generally depended on their education, prior experience and beliefs.

> For example, look at the emotional capacity. When you are working with that aspect, you use all the knowledge of the child that you have at that moment. But the treatment manual is not a pre-programmed instruction. (resp 1)

In addition to the perceived structure of RESTART-DCM treatment, a few therapists mentioned that their focus in therapy had shifted from decreasing the demands on the child to a combination of decreasing the demands and increasing the child's capacities. Language training in particular has become a more prominent goal for therapists.

With regard to the LP, in general, therapists were slightly more positive about the program than they had been before. They appreciated the clear structure and the perceived effect of the therapy within a relatively short period. They stated that the LP provided a very clear method for parents to work with. However, some therapists also reported the limited possibilities of the program if there is no, or minimal, progress in fluency in the first few months. They considered this was mainly related to the explicit goal of the LP as well as to the operant method (i.e., applying verbal contingencies to effectuate fluent speech). A lack of progress in the first months of the LP can also lead to decreased motivation on the part of parents, as mentioned by one participant.

> With Lidcombe, you are pretty much stuck with the contingencies. And, of course, at a certain moment, those contingencies lose their effect! When it takes longer. (resp 12)

> (...) I personally feel that I now have a better idea about the danger signals. Within the first three months, it should be working! And then you just have to be stricter, otherwise you do indeed lose the effect of the contingencies. (resp 2)

Three participants mentioned a positive side-effect of the LP, namely parents spontaneously praising their child in other situations than when their child speaks stutter-free.

> When the parents say something like, "I've been giving more compliments or have become nicer or more patient (…) in general in my parenting." (resp 6)

(…) That parents indeed say, "You have done that well!" But then about something completely different. That they are actually delivering contingencies for just about everything is quite funny, actually. (resp 2)

3. Attitudes toward success factors of the treatments

Key features of RESTART-DCM treatment which were mentioned as unique for this therapy are: slowing down the tempo of speech and communication, training in turn taking, and teaching parents to intentionally change aspects of their communication with their child (e.g., applying basic rules for communication and/or change specific aspects of communication that are supposed to trigger stuttering). The group discussion made clear that participants did not consider these factors as the only success factors; in their view there is no one factor which holds the key to success for RESTART-DCM treatment. One comment from the therapists:

Parents are also transformed into good communicators. (resp 2)

signals how participation in the process increases the communication skills of parents.

For the LP, a key factor for success was generally agreed to be the verbal contingencies, which are an essential part of the treatment [68]. Contingencies both after stutter-free speech as well as after unambiguous stuttering were considered to be necessary. Next to the contingencies, another factor considered as unique to the LP is the structure of transfer, especially in stage 2.

What I actually like, and what is really a part of Lidcombe, is stage 2. I often criticize it but….(resp 13)
What? (resp 10)
Stage 2. That it is structured and clear. Uh, that the line for the parents is short and simple: is it still going well, yes or no? That you are able to make a gradual transition and also have a good idea of when to sound the alarm. (resp 13)

This participant is referring to stage 2 of the LP where guidance is offered to the SLP for monitoring the process of fluency maintenance. Less stable fluency or a relapse could quickly be detected and result in taking a step back in the treatment process.

In addition to the focus on the unique therapy aspects, the moderator asked about similarities between the RESTART-DCM treatment and the LP. Seven similarities were mentioned (see also Table 7.2):

7

1) Both treatments start with one hour weekly therapy sessions;
2) Parents are trained by the SLP to practice with their child at home;
3) Parents practice with their child 15 minutes each day;
4) Overall, the intensity of time investment for parents is the same (although according to the group RESTART-DCM treatment requires a greater time investment of parents at the early stage of therapy, while the LP takes a longer follow-up);
5) Both treatments stress the importance of a pleasant interaction between child and parents. Within RESTART-DCM treatment it is stressed that a child should feel "heard and seen", especially during Parent-Child Interaction time, while the LP stresses the importance of the child's enjoyment when practicing fluent speech with the parent;
6) Both therapies aim to increase fluent speech;
7) Both treatment regimens require the SLP to constantly refine the treatment in order to enhance the child's fluency.

4. Learning outcomes and increased professionalism as a result of participating in the RESTART-trial

The majority of participants mentioned that, due to their participation in the RESTART-trial, they now begin their clinical contact with a comprehensive assessment of the child's speech and language, whereas beforehand they felt the urgency to start treatment as soon as possible (often because of parents' concerns and their request for therapy). The extensive assessment period enables the therapist to monitor the child's fluency over a few weeks. As already mentioned in the Introduction, in quite a few cases fluency spontaneously increased during the assessment period, thereby reducing the urgency to start treatment immediately (see the flow chart in[213]). Additionally, language assessment and video recordings of the parent-child interaction provide valuable diagnostic information which is helpful for therapists in their treatment decisions.

> Yes, I now easily start with a series of tests. During the first meeting I ask, "What do you want to know before leaving here today?" Because I think, maybe they already have an urgent question, and if so, I want to answer it. But I'm now more relaxed and patient about the rest of the process. (resp 9)
> About three contact moments, just to see if…(co-interviewer)
> Yes. And also to decide which method I will use. I think it is important to invest time in figuring that out now. (resp 9)

Two other learning outcomes of participating in the RESTART-trial were that (1) participants felt that working with the RESTART-DCM treatment manual that was designed for the trial had increased their expertise with regard to this treatment, and (2) participants were able to learn through interaction with each other during the regular meetings. The increased

expertise with RESTART-DCM treatment has already been addressed above (see theme *2. Changing beliefs about the potential of the treatments*). In addition, all participants recommended that, in order to provide high quality RESTART-DCM treatment, a 2-3 day course should be provided based on the RESTART-DCM manual (as is at present the case for the LP). Participants assumed this would be helpful not only for SLPs who start to treat preschool children who stutter, but also for SLPs who already work according a DCM approach. With regard to the latter learning outcome, the following two quotes clarify how the regular meetings provided both a safe environment to exchange experiences and the emergence of a shared basis for discussion:

> And discussing cases regularly on the basis of the treatment manual. That is, that you also have this shared basis from which you discuss cases. That also provided a real added value for me. (resp 2)

> (...) And what I personally think is great is that in our discussion meetings, that openness and safe feeling of exchanging experiences with colleagues – that this was possible and that the discussions were really useful - at least in my opinion. And that's really great. (resp 11)

DISCUSSION

This article reports the results of a focus group study that explored the attitudes, beliefs, and experiences of therapists participating in a RCT of stuttering therapy in preschool children who stutter. The results showed unequivocally that changes in attitudes and beliefs on RESTART-DCM treatment as well as on the LP had occurred due to their experiences in the trial. In particular, the potential that each treatment offered had become clearer over the course of the trial. This was mainly due to treating children as dictated by the randomization of treatments (whereby SLPs had to abandon their respective preferred treatment program), working intensively with both treatment manuals and meeting regularly to discuss and reflect on the interventions. The RESTART-DCM treatment manual which was developed at the start of the trial also contributed to enhanced insight into the opportunities of this treatment approach. Whereas prior to the trial nearly all SLPs had a preference for one of the treatments, in the final phase of the trial this preference no longer existed. Instead, all SLPs stated that their choice would now be determined by child-specific factors and the preference of the parents. A prevalent learning outcome for most participants in the RESTART-trial was the benefit of the comprehensive pre-therapy assessment. Not only did this result in a firm basis for therapy decisions, it also showed the relatively high percentage of children whose stuttering frequency decreased during this period and thus did not require immediate treatment. As a result of their participation in the

trial, participants shared the feeling that they were able to provide a higher quality of care because they were more flexible and more client-centered when considering treatment.

For example, participants in the focus group mentioned several child-related factors that they perceived to be important when considering a treatment. These factors included age, stuttering severity, capacities and temperament, speech- and language problems, and the child's speech attitude. The factors age and stuttering severity were included in the analyses of our trial data but they did not predict recovery with one of either treatment. However, this might also be due to a lack of statistical power to prove significance for subgroup analysis (see [213]). In the past decade several researchers have tried to unravel the complex relationship between stuttering and temperament (e.g., [100,101,238]), speech attitude [95,96] and/or speech- and language difficulties [31,32]. However, if and how these factors precisely interact, as well as their potential influence on treatment outcome, is still unclear. Future studies could investigate the impact of the factors perceived to be relevant for treatment outcome by the participants on long-term outcomes of the LP and RESTART-DCM treatment.

Three other findings related to the LP deserve attention. Firstly, several SLPs noticed that parents of children in the LP spontaneously began to praise their children outside the context of practicing fluent speech, in situations where a child showed other positive behavior (for instance, for putting away their toys). This observation might indicate a more general change in parent-child interaction during the course of the LP treatment. If and how this influences the child's fluency is beyond the scope of this study; however, it has been discussed in the literature that it is not yet known which components of the LP lead to successful therapy outcomes [207,209]. Treatment factors underlying both the RESTART-DCM treatment and the LP, such as an increase in self-efficacy because the child feels better "seen and heard" by the parent, might partly account for positive treatment outcomes. A second finding related specifically to the LP that was brought up by the focus group was uncertainty as to what action to take if progress within the LP slows down or ceases and the child is not yet speaking fluently. This difficulty is also acknowledged by the Lidcombe group and was put forward to argue the need for more evidence on the mechanisms of the LP by Hayhow [207]. Lastly, the SLPs noted the significance of stage 2 of the LP in reaching fluency and preventing relapse to occur. In contrast, an explorative study by Rousseau et al. [231] into the experiences of Australian SLPs with the LP revealed that about 30% of SLPs did not deliver therapy in stage 2 as it is intended to.

Interestingly, the shift from "therapeutic-centered care" (What suits me as a therapist?) to "client-centered care" (What suits the client?) shows the transition from the primarily perception-based practice which existed among Dutch SLPs to a more evidence-

based approach. Through participation in the trial, SLPs learned to look beyond their preconceptions and became willing to incorporate trial outcomes into their current practice. Since the participants consisted of pioneers within the field, this attitude shift can be of great value in implementing the results of the RESTART-trial, for instance in their lectures, courses, workshops, and other interactions with students. Together with this attitude shift, participation in the trial generally enhanced the clinical practice of participants and led to surprising learning outcomes with regard to the RESTART-DCM treatment, even though it has been the standard treatment in the Netherlands for the last decades. The authors share the strong belief of the participants that a course of several days should be made available for every SLP who starts working with this approach to better grasp the comprehensiveness and potential of this treatment, and thus ensure maximum benefit from the program. An explorative study by O'Brian et al. [232] showed significant effects of training on treatment outcomes for the LP. Education, whether it is a LP training or a DCM based treatment course, is best be followed by regular meetings to discuss specific issues related to treatment. If meetings are guided by a researcher with clinical experience, this type of training could lead not only to creating a degree of uniformity of clinical practice and enhancing professionalism in the field, but also to bridging the gap between research and practice by allowing opportunities for SLPs to extend their role to that of clinician-scientist.

At a more general level, the results of this focus group study demonstrate that qualitative research within the field of speech-language pathology can enhance the value of the more common quantitatively-orientated research. Where results of quantitative research provide insight into the *what* and *where* questions, reflection by SLPs on their current practice provides insight into the *why* and *how* questions such as "Why do therapists work in the way they do?' and 'How do therapists perceive different components of treatments?". Thus, it reveals (sub)conscious ideas underlying the professional activities of SLPs. Besides, a qualitative study can make the learning outcomes of participation in a clinical trial visible. Since there are few clinical trials in the field of speech- and language pathology, insight into these experiences are important for the interpretation of study results and the setting up of future trials.

The limitations of this study need to be considered. Firstly, the results presented in this study are limited to the attitudes of SLPs who participated in the RESTART-study. Since they had been actively working with both treatment manuals during the trial period, they were a source of rich information concerning both treatments and their comparison. However, it would be valuable to investigate if SLPs outside the RESTART-trial (or even outside the Netherlands) share similar attitudes on both treatments to the participants in our study. Comparing such attitudes would give insight into how far the evidence-based climate in the field of speech-language pathology (and stuttering therapy in particular) is currently

7

established. Another possible limitation of the study is the fact that there was a single meeting for the focus group discussion. An additional meeting, or in depth interviews with all participants, might have led to more detailed insight into attitudes of participants on the discussed topics. However, we feel confident that there was sufficient time for discussion as well as an open atmosphere in which SLPs could freely discuss their attitudes and beliefs about the two treatments. This was due to the familiarity of the setting and to the moderator and assistant interviewers being conscious of the importance of encouraging participants to express both positive and negative views.

In conclusion, this study showed evidence of the evolution in attitudes and beliefs of Dutch SLPs with regard to two widely applied therapy programs for young children who stutter: RESTART-DCM treatment and the LP. It showed specifically that the SLPs are open to an evidence-based approach in the field of speech and language therapy. This work increases our understanding of how attitudes and beliefs of therapists play a role in the uptake and utilization of therapies and demonstrates the benefit of qualitative research. The results and recommendations should prove of value both in implementing the RESTART-trial results and in training of SLPs.

ACKNOWLEDGEMENTS

8

Discussion

INTRODUCTION

Prior to the start of this thesis, the body of evidence to support clinical decision making by speech-language pathologists (SLPs) was limited. Rigorous clinical studies were scarce and little attention was given to evidence based practice (EBP) in speech-language pathology. Accordingly, decision making by SLPs was primarily opinion- and experience-based. The field of developmental stuttering was no exception. In the past decade, the health care system has been changed drastically. Rising medical costs have induced an increasing emphasis on *value for money* in reimbursement decisions and, consequently, on the empirical validation of treatments. Reimbursement of speech-language treatment has become under pressure. The Dutch government introduced a plan to cut down reimbursement from the year 2015, but to date these plans have not been implemented. Meanwhile, EBP has gradually emerged within the field of speech-language pathology, as for instance reflected by the number of publications on this topic. A PubMed search revealed that, of a total of 74 scientific publications on EBP and communication disorders, 20 have been published before and 54 since the year 2007.[1] Results of the first phase III randomized trial into stuttering treatment in children have been published in 2005 [6], and in 2014 a guideline for the treatment of developmental stuttering was published [65]. It was the first mono-disciplinary guideline for speech and language disorders in the Netherlands, however, most of its recommendations were still based on limited evidence.

This thesis addressed two main gaps in the evidence base for stuttering treatment. First, outcome studies have traditionally focused on stuttering symptoms. Broad, patient reported, outcome measures like quality of life have not often been applied, while such instruments provide essential information on the impact of the disorder on daily living; information that is increasingly demanded by policy makers. Second, while it is generally agreed that developmental stuttering should be treated in the preschool years [11,184], little is known about the effectiveness of available treatments for young children who stutter. A systematic review and meta-analysis by Nye et al. [66] concluded that, in the limited data available, the Lidcombe Program offers the best empirically supported intervention, and that there is insufficient data using high research standards to support the effectiveness of other treatment approaches. The main objective of this thesis was to fill these two gaps by studying the impact of stuttering and the effectiveness and cost-effectiveness of treatment in preschool children.

---

1   Search string: evidence-based practice [Title/Abstract] AND (speech language [Title/Abstract] OR stuttering [Title/Abstract]).

This chapter discusses the main findings of this thesis and the relevance for the evidence base underlying stuttering treatment decisions. Furthermore, it describes methodological issues and implications for clinical practice and policy making. The chapter concludes with recommendations for future research.

MAIN FINDINGS

In the study presented in chapter 2 we found that stuttering did not affect the health-related quality of life (HRQOL) of children in the preschool years. However, in our study into the HRQOL of adults who stutter (chapter 3) we found that severe persistent stuttering was associated with a reduced HRQOL. The impact of mild and moderate stuttering appeared to be limited. Coping style appeared to be related to HRQOL in almost equal amounts as stuttering severity. One of the instruments applied in this study was the Dutch OASES for adults (OASES-A-D), which is a questionnaire to comprehensively measure the impact of stuttering on a person's life. The psychometric properties of the OASES-A-D were assessed in a separate study, of which the results are presented in chapter 4. The OASES-A-D was found to be a reliable and valid tool to assess the impact of the stuttering disorder on daily living (chapter 4).

The current standard treatment for preschool children who stutter in the Netherlands is treatment based on the Demands and Capacities model (DCM), hereafter referred to as RESTART-DCM treatment. In the RESTART-trial, we showed that the Lidcombe Program (LP) is an effective and cost-effective alternative to RESTART-DCM treatment (chapter 5 and 6). The LP decreased stuttering frequency more quickly during the first three months of treatment. At 18 months, clinical outcomes were comparable, though most outcome measures were slightly in favor of the LP. In both treatment groups, stuttering frequency hardly further diminished after three months, while about 30% of children was still on treatment at 18 months. Costs for one additional child who stopped stuttering with the LP as well as extra costs per QALY, were favorable for the LP compared to RESTART-DCM treatment. Participating in a randomized trial appeared to affect the process of individual clinical decision making, as discussed in chapter 7.

In the next part I will elaborate on these findings by discussing three themes that emerged from this thesis. These themes addressed the following topics: (1) when to initiate treatment; (2) which treatment to choose; (3) how to improve the therapy process.

STRENGTHENING THE EVIDENCE BASE FOR STUTTERING TREATMENT

Early intervention for developmental stuttering

It is current standard practice to treat children who stutter in the preschool years [11,184]. Arguments for early intervention are primarily based on studies showing that the chance of recovery decreases with time since stuttering onset [34,44]. Our results support the need for early intervention, as persistent severe stuttering was associated with a significant impact on HRQOL. Treating children in the preschool years reduces the likelihood that stuttering becomes persistent and has a negatively impact on daily living.

Despite the apparent need for early intervention, instant treatment for every child that stutters who presents to the clinic is usually not required. A relatively large number of children recover spontaneously, and delaying treatment for one year is not associated with a worse treatment outcome [239,240]. Besides, the results presented in chapter 2 suggest that the well-being of children is hardly affected at preschool age. This suggests that treatment may well be preceded by a period in which the child's stuttering is solely monitored. Treatment would then only be required if natural recovery failed to occur. Monitoring a stuttering child in the first year after onset is also recommended in the current guideline [65]. Obviously, high parental concerns or a clear burden on the child in the first year after onset could be reasons to start treatment earlier.

Findings presented in this thesis suggest that Dutch SLPs are generally inclined to start treatment earlier than the proposed waiting time of one year. The slow inclusion of children in the RESTART-study was for a great part related to a relatively high percentage of children whose fluency increased spontaneously during the comprehensive assessment phase at baseline, thereby reducing the urgency to start treatment (chapter 5). The focus group discussion that was held in the context of the RESTART-trial (chapter 7) confirmed that most SLPs were used to start treatment as soon as possible before participating in the trial, often because of parents' concerns and their request for therapy. In the final phase of the trial, most participating SLPs had integrated the comprehensive assessment of the child's speech and language in their routine practice. This enabled them to make more informed decisions on the start and choice of treatment.

Selecting a relevant treatment

Currently, most Dutch children are treated according to a DCM based approach. This thesis showed that the LP is a good alternative to DCM based treatment. Children in the LP group even improved slightly more in fluency as well as on most secondary outcomes, although the differences did not reach statistical significance (chapter 5). Mean total costs per child

were only marginally higher for the LP group (chapter 6). The mean difference of €168 may diminish in the long term, as slightly more children in the RESTART-DCM group were on treatment at 18 months. Although our results do not conclusively prove that the LP and RESTART-DCM treatment are equivalent options, there is sufficient evidence to support the implementation of the LP in current Dutch practice. If the LP would be available on a large scale, there will be freedom of choice to decide on what works best for an individual child and the family, which in turn creates opportunities to tailor therapy to individual needs.

Since the results do not indicate a clear preference for the LP or RESTART-DCM treatment, we concluded that shared decision making by SLP and parent is recommended. The results of the focus group meeting showed that, in the final phase of the trial, participating SLPs gave a great weight to child-specific factors and the preference of the parents in selecting a treatment (chapter 7). Factors that may be important for parents are related to type of treatment, parent's abilities, duration and costs. The result that the LP decreased stuttering more rapidly during the first three months might be an argument for parents to prefer the LP. We incorporated the child-specific factors age, stuttering severity, and time since onset in our analyses in chapter 5, but none of them were related to differences in treatment outcome with the LP or RESTART-DCM treatment. Inspection of the data, however, showed a possible effect of severity, with a higher percentage of children who initially stuttered severe classified as non-stuttering at 18 months in the RESTART-DCM group. The lack of statistical significance for this analysis might be due to the small groups resulting of splitting up severity groups. Thus, our results so far do not enable us to provide recommendations on whether particular child-specific factors are relevant with regard to decisions on selecting the LP or RESTART-DCM treatment.

In sum, our findings do not favor either the LP or RESTART-DCM treatment for preschool children who stutter and support a shared decision making by parents and clinician in the process of selecting a treatment.

Improving the therapy process

The RESTART-trial revealed two other findings relevant for future improvement of stuttering treatment in preschool children.

The first result is related to the professional standard of DCM based treatment. The focus group meeting revealed that, before the RESTART-trial, and accordingly before SLPs started working with the RESTART-DCM manual, there was a perceived lack of guidance to take decisions on which components of the comprehensive framework to choose for an individual child. The RESTART-DCM treatment manual appeared highly instrumental in providing structure to the treatment. Working with this manual led to an increased insight into the comprehensiveness and potential of this treatment, and accordingly to a higher

level of expertise. Of note is that the group of participating SLPs in the RESTART-trial consisted of clinicians who had many years of clinical experience with DCM based treatment. This led us to conclude that, although DCM based treatment is the current Dutch standard for treating preschool children who stutter, there is considerable potential to increase the professional level of SLPs outside the RESTART-trial working with this approach.

The second finding refers to the data on stuttering frequency presented in chapter 5, which suggests that cost-effectiveness of treatment might be improved by reducing treatment duration and/or intensity. The largest speech improvement was obtained in the first three months after treatment onset. There was only a very gradual decrease in stuttering frequency in the following 15 months, while about 30% of children was still on treatment at 18 months. The merits of extending treatment beyond three months are doubtful, especially if one considers that the impact of mild and moderate stuttering on quality of life is negligible. In Dutch daily practice, the reimbursed number of stuttering treatment sessions is not restricted by health care policy. Criteria to terminate treatment are generally related to the objective of reaching fluent speech. For RESTART-DCM treatment, active monitoring instead of weekly treatment sessions during the very gradual improvement phase may be sufficient to prevent relapse and effectuate the gradual normalization of speech fluency. Regarding the LP, maintenance treatment is an important aspect of behavioral treatment, in order to reduce the risks of a relapse. However, there may be potential to reduce the length or intensity of the maintenance phase. A study by Rousseau et al. [231] showed that about 30% of children treated with the LP in Australia did not participate in Stage 2 (maintenance phase) at all.

Relevant in this regard is the topic of how to define normally fluent speech that is usually strived for. This is an ongoing topic of debate in literature [241-243]. The mean %SS in both treatment groups in our study had already dropped to about 3% SS after three months of treatment, which has been suggested to be the upper limit for normally fluent speech. Non-stuttering at 18 months was defined as less or equal than 1.5% SS, based on a recent study by Clark et al. [191] showing a mean frequency of 1.5% SS in CWNS. The LP manual that was applied in our study [48] specifies a slightly stricter goal of treatment: no stuttering, defined as a frequency less or equal than 1% SS, which must be maintained for a long time. As a result of this discrepancy, over 20% of children in the LP group classified as not stuttering anymore was still on treatment at 18 months post-treatment onset. Although criteria to withdraw and terminate RESTART-DCM treatment are less strictly defined, an approximately similar percentage of children classified in our study as not stuttering was still on treatment at 18 months. So, explicit or implicit criteria on fluency applied in stuttering treatment might be too stringent in light of the study by Clark and colleagues.

8

Considering also the possibility that recovery toward fluency -in whatever way defined- will not stop if therapy intensity is reduced (it may just take maturation time for the very mild stuttering to completely disappear), reconsidering the criteria to terminate treatment seems warranted.

METHODOLOGICAL CONSIDERATIONS

Methodological considerations related to the RESTART-trial

Designing and conducting an RCT with a piggyback economic evaluation brings along several challenges and, accordingly, choices to be made.

One challenge we faced was related to defining and measuring stuttering, which is, as said above, an ongoing topic of debate in literature [241-244]. Our primary outcome measure was based on the calculation of the mean stuttering frequency in daily conversational samples by expert raters. This measure does not take into account non-observable stuttering behaviors like facial movements, nor the duration of the stuttering events. Nevertheless, since it is known that stuttering in young children mainly constitutes of observable disruptions in the flow of speech (i.e., repetitions of sounds and syllables, prolongations, and/or blockages), stuttering frequency is the most widely applied outcome measure in research into childhood stuttering. Another choice regarding the primary outcome was the cut-off score. We chose a cut-off score of 1.5% SS [191], but, as said above, there is a grey area between about 1% and 3% SS. Applying a cut-off score of ≤ 1% SS or ≤ 2% SS did, however, not affect our results (chapter 5). This strengthens our finding that both treatments are comparably effective after 18 months of therapy.

In designing a trial with a piggy back economic evaluation, the concepts of internal and external validity comes into play. Whereas an RCT generally aims for a high internal validity in order to establish treatment *efficacy* (i.e., the extent to which an intervention is beneficial under ideal conditions), a cost-*effectiveness* study desires a high external validity. Outcomes obtained under conditions close to daily practice have the potential to assist clinicians, clients and policy makers in making informed decisions that will improve health care at both the individual and the population level. On the efficacy-effectiveness continuum, our study-design was more toward the effectiveness-end. Sufficient internal validity was obtained by, among others, the process of randomization, the use of the RESTART-DCM manual, and regular contact with participating SLPs to enhance treatment fidelity. We achieved high external validity by, among others, sampling procedures to enroll participants that were representative of the clinical population, and SLPs offering the treatment as they would do in daily practice. Despite these achievements of external

validity, the effect of participation in a trial for SLPs and parents on the generalizability of our outcomes cannot be totally ruled out. For instance, treatment adherence rates might have been higher than in daily practice. In addition, participating SLPs probably differs from other SLPs in some respect, e.g., with regard to their level of experience and their attitude toward letting go of preferences for either treatment option. Furthermore, chapter 7 showed that SLPs became more proficient in clinical decision making and delivering treatment during the trial. It remains unclear whether this might have influenced the results; possibly the obtained percentages of children who did not stutter anymore at 18 months in our study are slightly higher than in daily practice.

Methodological considerations related to quantifying the impact of stuttering on daily living

Generic preference-based HRQOL instruments can be used to generate QALY estimates and thus to compare the impact of stuttering to that of other conditions on the same scale. However, the use of these instruments requires addressing several issues, most of which are related to the concept of validity. In chapter 2 and 3 we already addressed some of these issues, like the differences in results obtained by different HRQOL measures. This section will discuss three other, more generic, validity issues.

A first issue in the context of validity of HRQOL instruments is that these instruments need to be sensitive and specific enough to pick up small differences that are of importance to people living with the condition [245], in this case stuttering. It is questionable if generic HRQOL measures incorporate relevant domains that sufficiently capture the social and psychological well-being effects associated with stuttering. The impact of stuttering may be best assessed by condition-specific HRQOL measures, yet this kind of instruments do not exist. The OASES-A-D incorporates one section on quality of life and, due to the broad range of questions, provides a comprehensive assessment of the impact of stuttering on a person's life. This instrument is therefore relevant for use in clinical practice, but it cannot be used to capture the benefits of stuttering interventions in terms of QALYs.

A second validity issue relates to the measurement of HRQOL in children, which is generally acknowledged to be complex. Several problems have been raised in the literature. Briefly discussed, they relate to (1) a lack of consensus on the conceptual definition of HRQOL as it relates to children [246]; (2) the proxy issue [247]; and (3) the suitability of HRQOL measures adapted from instruments initially developed for adult populations [248,249]. With regard to the latter, the health state descriptions of the HUI3, applied in the study in chapter 2, may not fully reflect the health dimensions applicable to young children [245].

8

Thus, to adequately measure the effect of stuttering (treatment) on HRQOL by preference-based measures, we are in need of an instrument that includes relevant domains, can detect a small change in health state, and can be applied in young children. At present, an instrument possessing all these characteristics seems not to exist. Since the HUI3 includes domains on speech and on psycho-social aspects, and since this instrument showed some variability in our study, the HUI3 might be the current best available alternative instrument in this regard [151,245].

Nevertheless -and this is the third issue related to validity- it is open to discussion whether preference-based generic HRQOL measures could be capable of determining the impact (or value) of stuttering at all. Since people who stutter can face many hurdles in verbal communication throughout the day, stuttering may well be a "high attention grabber" for an individual who stutters. As proposed by Paul Dolan, adaptation - the process of adjustment to new or changed circumstances - to a condition that potentially has such a high impact on the frequency, intensity and duration of one's thoughts and feelings is rather difficult [148,250]. Adaptation is probably further hampered by the variable character of the stuttering severity. This is probably quite hard to estimate for the general public, whose preferences are used to express the relative desirability -value- of different health states in generic preference-based HRQOL measures as the HUI3. As a result, the general public might overestimate the HRQOL of people who stutter. However, the potential influence of stigma on preference-based generic HRQOL values might work out the other way around. That is, public stigma related to stuttering (e.g., in the form of negative stereotypes, prejudice, and discrimination [251,252]) might degrade HRQOL values for stuttering given by the general public.

An appealing alternative in this regard might be to use "patient preferences". However, it has been argued that preference-based valuations obtained in patients suffer from the same problem as those obtained in the general public: the values elicited reflect imaginations about the impact of a health state when people are focusing attention on the impact [253]. Thus, they do not reflect the (future) value while experiencing a particular health state. Dolan [253] argues that these values are therefore not useful in establishing "how severe different conditions are when they drift in and out of attention in the day-to-day experiences of life" ([253], p.2). Dolan claims that we should seek for more direct measures of the value associated with different health states, like directly measuring happiness [250,253]. It is conceivable that a high attention-seeking condition like stuttering will have a greater impact on happiness than, for instance, suffering from some problems walking about. This new avenue might therefore be of great relevance to the field of stuttering.

IMPLICATIONS

Clinical implications

At present, most Dutch SLPs are trained and experienced in DCM based treatment, which forms a part of their regular education. As discussed above, there is sufficient evidence to support the uptake and utilization of the LP in current Dutch practice. In particular, SLPs who currently only provide DCM based treatment should be encouraged also to become a LP certified clinician by following the Lidcombe Program training. The LP should also receive considerable attention in regular educational programs.

With regard to DCM based treatment, SLPs should be encouraged to apply the RESTART-DCM manual. This is expected to lead to more uniformity and transparency in the goals, content and principles of the treatment. This, in turn, can facilitate communication and negotiations with other players in the field, like other health care providers and decision makers. Developing a course based on the RESTART-DCM manual for SLPs who starts working with this treatment is highly recommended.

This thesis brought about several aspects of clinical decision making in stuttering that must also be faced by the professional community in the field of speech-language pathology. First, EBP prescribes incorporation of preferences and values of patients in treatment decisions. Based on our result we strongly propose shared decision making by SLP and parents in the process of selecting a relevant treatment for an individual child. This is in line with the recently published Dutch guideline on developmental stuttering [65]. Second, in light of arguments of cost-effectiveness, criteria for initiation and termination of treatment deserve more attention of experts in the field of stuttering. These aspects have not yet gained much attention from clinicians. In order to address them, it is vital that SLPs acknowledge the relevance of research in this area.

Hopefully this thesis has not only contributed to a growing evidence base, but also toward helping SLPs to develop an evidence-based mind-set. In this regard, it is highly recommended that SLPs who participated in the RESTART-trial will be involved in the implementation of our study results and in the initiation of further research. Not only does this group of SLPs consist of "early adopters" [233] who can inspire change among the majority of SLPs, chapter 7 also revealed that they had made a transition from a primarily opinion-based practice to a more evidence-based approach as a result of their participation in the trial. As such, they prove to be a valuable example in EBP for colleagues. If SLPs can achieve not only to provide excellent care but also to improve the efficiency of care, they will have a competitive position in negotiations with insurance companies and health care policy makers.

8

With regard to treatment for adults who stutter, our results suggest that assessment of stuttering in adult clients should explicitly address potential quality of life impairments as well as coping style as one of its determinants. The importance of a multidimensional approach addressing quality of life has already been acknowledged by the guideline [65]. Although further research on test-retest reliability and potential improvements for decreasing the length of the OASES-A-D would be useful, we consider the OASES-A-D a valuable tool for comprehensively assessing the impact of the stuttering disorder in clinical practice. The relevance of addressing coping style, which was equally related to HRQOL as stuttering severity, should be more clearly articulated in the guideline. Treatment aimed at decreasing stuttering severity and increasing adequately coping strategies might possibly result in the best chances for enhancement of a person's quality of life.

Implications for policy makers

Given the small differences in effects and costs, we propose that the LP and RESTART-DCM treatment both remain in the basic health insurance package. Supporting the implementation of the LP and training SLPs in the RESTART-DCM method is associated with additional costs, but could eventually lead to better quality of care and potential reductions in health care expenditures, since SLPs mastering both type of treatments are more flexible in selecting the appropriate treatment and adjusting the treatment according to individual needs.

Cost-effectiveness alone is never sufficient for rational decision making on reimbursement of treatments. In particular, the burden of a disease is known to interact with cost-effectiveness considerations. The higher the burden of disease, the more willing society is to accept a poor cost-effectiveness. This can be shown by the example of Viagra, which is known to be highly cost-effective but is not collectively reimbursed [254,255]. Lung-transplantation, on the other hand, is known for its unfavorable cost-effectiveness, yet the reimbursement is no matter of debate [256]. Our studies showed a high HRQOL, and thus a low burden of disease, in children who stutter in the preschool years, but a higher burden of disease for adults with persistent stuttering. Persistent stuttering has also shown to be related to substantial health care costs [212], while chapter 6 of this thesis showed that the absolute costs of stuttering treatment in the preschool years were relatively low (about €3000 in 18 months). Therefore, we strongly recommend full reimbursement of stuttering treatment in the preschool years. In case parents should pay treatment for their stuttering preschooler by themselves, they possibly will wait too long to seek treatment, thereby increasing the chance that stuttering will be more difficult to treat. On the other hand, because of the high chance of spontaneous recovery in the early years [12,34], treating many children that would recover spontaneously anyway is neither regarded a cost-effective strategy. Encouraging SLPs to follow the guideline by Pertijs et al. [65] in the process of clinical

decision making regarding treatment onset could be a cost-effective way to promote effective and efficient healthcare [257].

The relatively high disease burden in terms of HRQOL associated with severe stuttering in adulthood yields also an argument in favor of reimbursing (effective) treatments for this population. Disease burden associated with mild and moderate stuttering appeared to be low, however a limitation of our study is that we only applied subjective stuttering severity measures. Furthermore, other criteria that were outside the scope of this thesis (e.g., treatment effectiveness and cost-effectiveness) are of course also relevant in reimbursement decisions on treatment for adults who stutter.

RECOMMENDATIONS FOR FUTURE RESEARCH

Besides providing valuable new insights into the impact and treatment of stuttering, this thesis offers several avenues for future research.

First, the very slight differences in outcomes between the LP and RESTART-DCM treatment suggest to investigate the underlying factors that contribute to a positive treatment outcome. Further research could explore if prognostic indicators of spontaneous recovery, like a shorter time since onset or recovery of stuttering in the family, or aspects related to the child's speech-language profile, can also predict treatment outcome– regardless from the provided therapy. Exploring treatment agents, including common factors, underlying the LP and RESTART-DCM treatment is another interesting research area in this regard. Especially for those children who are likely to recover spontaneously, common factors may be enough to accelerate the process of recovery. Since there is still much to be learned on the nature of the stuttering disorder and the role of treatment in recovery, this kind of data could be of great value.

Second, further research should address the efficiency of treatment by determining whether treatment duration or intensity can be reduced, without reducing treatment effectiveness. This avenue of future research could result in a relatively large reduction of treatment costs.

Third, the need for early intervention could be justified by research on the long term effects of treatment in the preschool years. Besides following the children that participated in the RESTART-study, which is planned to be initiated, a model-based study incorporating data from the RESTART-trial and HRQOL data of adults who stutter could quantify the gains of early intervention in terms of prevented HRQOL losses associated with persistent stuttering.

8

CONCLUDING REMARKS

The overarching goal of every discipline in health care is maximizing value for patients. How to define *value* is yet a topic of continual debate. Ask a patient, a health care provider, a policy maker and a health economic researcher, and you will probably get four different answers. This thesis aimed to bridge the gap between different stakeholders in the field of developmental stuttering by strengthening the evidence base and addressing various aspects of the "value" of stuttering and its treatment. There are still steps to be taken in order to improve the care of people who stutter. This is only possible through active collaboration between different stakeholders in the field. Clinicians, researchers, policy makers, people who stutter and their families can help this process along- and all will benefit from doing so.
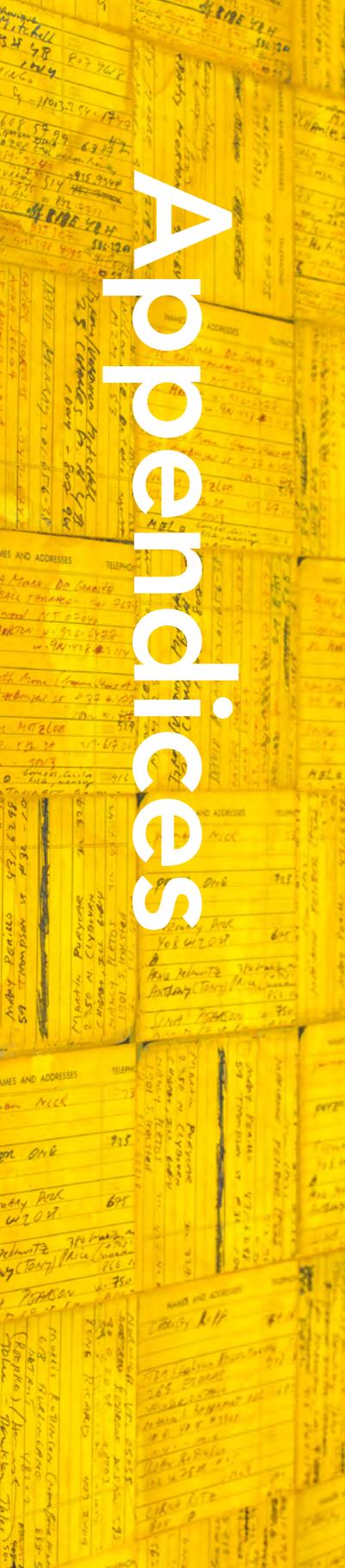
Appendices

Summary

Samenvatting

List of abbreviations

PhD portfolio

List of publications

About the author

Dankwoord

References

SUMMARY

Stuttering is a developmental disorder characterized by an abnormally high frequency of disruptions in the flow of speech. Stuttering onset is most likely in children between three and five years of age. About 5 to 11% of all children stutter at some time in their life. Most children outgrow the disorder of stuttering before the age of 10 years. Although several factors related to a higher chance for recovery have been identified in the past decennia, it cannot be predicted whether a child will recover naturally. The chance for complete recovery diminishes as children grow older, and persistent stuttering is assumed to negatively impact on a person's daily life. Therefore, stuttering is generally recommended to start in the preschool years (before the age of six years).

This thesis is devoted to studying the impact of developmental stuttering in childhood and adulthood, and the outcomes of treatment in the preschool years. Chapter 1 introduces and motivates this thesis. It describes that the introduction of evidence based practice (EBP) in the paramedical field led in the nineties to a growing awareness among speech- and language pathologists (SLPs) to incorporate scientific evidence in their clinical decisions. However, a decade ago high quality evidence to support treatment decisions was still scarce. Policy changes in healthcare and overall increasing health care expenditures exposed the urgency of research on effects as well as on costs of alternative treatment strategies. The field of developmental stuttering was no exception with respect to the poor evidence base underlying clinical decision making, as explained in chapter 1. It is substantiated that there is a need for data on the disease burden associated with stuttering, as well as high standard research into the effectiveness and cost-effectiveness of treatment in the preschool years.

Chapter 2 to 4 of this thesis focus on the impact of stuttering on daily living. Chapter 2 explores the burden of stuttering in terms of health-related quality of life (HRQOL) among preschool children. Parents of children participating in the RESTART-study, a randomized controlled trial into stuttering treatment in preschoolers, filled in several questionnaires on the HRQOL of their children. Data collected before the start of treatment were compared to HRQOL data of a reference population of Dutch children who did not stutter. Results indicate that the HRQOL of preschool stuttering children is comparable to that of non-stuttering children. Furthermore, stuttering severity and time since onset of stuttering were not related to HRQOL scores. It was concluded that, generally speaking, the HRQOL of preschool children who stutter is not diminished. However, only a small number of children with severe stuttering participated in this study. Therefore, we recommend future research including a larger cohort of children with severe stuttering. In a future study, it would also be relevant to study the course of HRQOL over time and incorporate additional parameters such as characteristics of the child and his environment.

S

Chapter 3 shows that, if stuttering persists into adulthood, the impact on HRQOL becomes more prominent. In the study presented in this chapter, 91 adults who stutter participated. They were partly recruited informally, and partly from stuttering and/or speech- and language therapy centers throughout the Netherlands. By a comprehensive approach their stuttering severity, functioning, HRQOL and coping style was investigated. Particularly severe stuttering was found to be related to a lower HRQOL. Not only the speech domain of one of the HRQOL instruments (HUI3) showed lower values for higher stuttering severity, but also the emotional and social domains. On two instruments (EQ-VAS and OASES-A-D), the scores of adults in therapy indicate a significant higher impact on daily living than the scores of adults not in therapy. Results show that coping style was equally strongly related to HRQOL as was stuttering severity. In particular, higher scores on the emotion-oriented subscale of the CISS coping instrument (reflecting a more negative way of dealing with emotions) were correlated with lower HRQOL values.

The studies described in chapter 2 and 3 applied generic HRQOL questionnaires. That is, they measured HRQOL in terms that are relevant to everyone. Generic instruments can therefore be applied in all kind of populations. Disease specific instruments, on the other hand, include only domains that are deemed to be relevant to people with the disorder under study. Preferably, a combination of generic and disease specific instruments is applied in the assessment of HRQOL. However, no Dutch validated stuttering specific HRQOL instrument was available at the time we conducted our studies. A few years before starting our study, the English OASES for adults was developed and psychometrical validated. The OASES is a broad instrument, including a section on the measurement of quality of life. It provides information on the experience of the stuttering disorder from the perspective of the adult who stutter himself. Chapter 4 described the translation of the OASES into Dutch and the examination of the psychometric properties of this instrument. The Dutch OASES for Adults (OASES-A-D) was found to be a reliable and valid instrument for providing a comprehensive assessment of how stuttering affects the lives of individuals who stutter. All sections of the OASES-A-D were able to differentiate between groups of participants with different levels of stuttering severity. We concluded that the OASES-A-D could be valuable in clinical practice to assess the impact of stuttering on daily living, although further research on the test-retest reliability and potential improvements for decreasing the length of the instrument is recommended.

Chapter 5 to 7 of this thesis were devoted to the RESTART trial (Rotterdam Evaluation study of Stuttering Therapy in preschool children- A Randomized Trial). In this 18-month trial, 199 preschool children who stuttered for at least six months were randomized to the Lidcombe Program (LP) or treatment based on the Demands and Capacities Model (RESTART-DCM based treatment). The latter has been the Dutch standard treatment for preschool children

who stutter since the eighties. It is an *indirect* approach, in the sense that treatment focuses on manipulating child related and environmental factors assumed to influence the child's speech fluency. In contrast, the LP is a *direct* approach that targets the child's speech fluency directly by means of behavioural modification (i.e., treatment based on operant conditioning). The interventions were offered by 24 SLPs in private practices throughout the Netherlands. Health outcomes and costs were evaluated at baseline and 3, 6, 12, and 18 months after the start of treatment.

Chapter 5 describes the results of the comparative effectiveness of the LP and RESTART-DCM treatment. The primary outcome measure was the percentage of children who did not stutter at 18 months. Secondary outcome measures included stuttering frequency, stuttering severity ratings by the parents, therapist, and child, HRQOL, emotional and behavioral problems, and speech attitude. Treatment by the LP decreased stuttering more quickly than RESTART-DCM treatment during the first three months of treatment. At 18 months, however, clinical outcomes for both treatments were comparable. Slightly more children in the LP group were defined as non-stuttering compared to the RESTART-DCM group (76.5 and 71.4%, respectively) but this was a non-significant difference. We concluded that both treatments are roughly equal in treating developmental stuttering in ways that surpasses expectations of natural recovery. An interesting finding was that, for both treatment groups, the largest speech improvement was obtained in the first three months after treatment onset and that there was only a very gradual decrease in stuttering frequency in the following 15 months, while about 30% of children was still on treatment at 18 months.

Chapter 6 reports on the results of an economic evaluation that was conducted alongside the RESTART-trial. The total 18-month costs were related to the number needed to treat (NNT; the number of children who need to be treated with the LP in order to have one extra child defined as non-stuttering at 18 months post-treatment onset, compared to RESTART-DCM treatment), and to quality adjusted life years (QALYs) based on the measurement of HRQOL. Since HRQOL values were high at baseline, only slight gains in quality of life were established during the follow-up. At 18 months, health outcomes were slightly better in the LP group, but differences were statistically and/or clinically not significant. Total costs were somewhat higher for the LP group. The difference in costs could possibly diminish in the long term, as slightly more children in the RESTART-DCM group were on treatment at 18 months. Based on the NNT of 20 and the difference in total costs of €168, the estimate of the incremental costs per NNT was €3360. Thus, an extra investment of €3360 would result in one more child classified as non-stuttering with the LP compared to RESTART-DCM treatment. The costs to obtain extra improvement in HRQOL by the LP compared to RESTART-DCM treatment were €10413 per extra V-QALY and €18617 per extra U-QALY,

which indicates good cost-effectiveness of the LP. Based on the results presented in chapter 5 and 6, it is concluded that LP is a good alternative to RESTART-DCM treatment in Dutch primary care.

Decisions on treatment selection and implementation of study results are known to be influenced by ideas, perceptions and experiences of clinicians. Chapter 7 therefore describes the results of a focus group meeting, in which participating SLPs discussed their attitudes and beliefs with regard to the LP and RESTART-DCM treatment, and how these might have changed as a result of participating in the RESTART-trial. The results show that, in the final stage of the trial, participating SLPs found themselves to be more flexible and more client-centered when considering treatment. The potential that each treatment offered had become clearer over the course of the trial. Whereas prior to the trial nearly all SLPs had a preference for one of the treatments, in the final phase of the trial this preference had diminished. Instead, all SLPs stated that their treatment selection would now be determined by child-specific factors and the preference of the parents.

Chapter 8 discusses three themes that emerged from this thesis, related to (1) early intervention; (2) selecting a treatment, and (3) improving the therapy process. First, as persistent severe stuttering in adulthood was associated with a significant impact on HRQOL, this supports the need for early intervention. The HRQOL of preschool children who stuttered was found to be hardly affected, indicating that treatment may well be preceded by a period in which the child's stuttering is solely monitored, as also recommended in the current guideline on stuttering. Second, our findings do not favor either the LP or RESTART-DCM treatment for preschool children who stutter and support a shared decision making by parents and clinician in the process of selecting a treatment. Third, the therapy process can be improved by encouraging the use of the RESTART-DCM manual by SLPs who want to treat a child according to a DCM method, as well as by reconsidering the merits of treatment beyond three months. Methodological issues related to the RESTART-trial and to quantifying the impact of stuttering on daily living are also discussed in chapter 8. In particular, validity issues regarding the use of HRQOL measures to assess the impact of stuttering are considered. The chapter concludes with implications for clinicians and policy makers, and recommendations for future research. There is sufficient evidence to support the uptake and utilization of the LP in current Dutch practice. Given the small differences in effects and costs, we propose that the LP and RESTART-DCM treatment both remain in the basic health insurance package. Future research should investigate the underlying factors that contribute to a positive treatment outcome, whether treatment duration or intensity can be reduced without reducing treatment effectiveness, and the long term effects of treatment in the preschool years.

SAMENVATTING

Stotteren is een ontwikkelingsstoornis die gekenmerkt wordt door onderbrekingen in de vloeiendheid van de spraak. Stotteren ontstaat meestal tussen het derde en vijfde levensjaar. Ongeveer 5 tot 11% van alle jonge kinderen stottert een bepaalde periode. De meeste kinderen groeien er overheen voor de leeftijd van 10 jaar. Ondanks dat er factoren bekend zijn die de kans op herstel van stotteren vergroten, is het niet te voorspellen of een kind over het stotteren heen zal groeien. De kans op volledig herstel neemt af met de leeftijd. Blijvend stotteren wordt in verband gebracht met negatieve gevolgen voor het dagelijks leven. Het wordt daarom aangeraden om behandeling voor stotteren te starten in de voorschoolse leeftijd, dat wil zeggen voor de leeftijd van zes jaar.

Dit proefschrift richt zich op het onderzoeken van de impact van ontwikkelingsstotteren op de voorschoolse en volwassen leeftijd en op de uitkomsten van behandeling in de voorschoolse leeftijd. Hoofdstuk 1 is een inleidend hoofdstuk, waarin de motivatie voor het onderzoek uiteen wordt gezet. De introductie van evidence based practice (EBP: op bewijs gebaseerd handelen door een therapeut) in de paramedische sector in de jaren negentig leidde tot een vergroot bewustzijn onder logopedisten om wetenschappelijk bewijs te integreren in het klinisch handelen. Desondanks was er begin deze eeuw nauwelijks bewijs van hoog wetenschappelijk niveau voorhanden. Veranderingen in het beleid en stijgende kosten van de gezondheidszorg leidden verder tot een toenemende vraag naar op bewijs gebaseerde zorgverlening. In hoofdstuk 1 wordt uitgelegd waarom er enerzijds meer gegevens nodig zijn over de invloed die stotteren heeft op het dagelijks leven en anderzijds meer gedegen onderzoek nodig is naar de effectiviteit en kosteneffectiviteit van behandeling in de voorschoolse leeftijd.

De hoofdstukken 2, 3 en 4 hebben betrekking op de impact van stotteren op het dagelijks leven van personen die stotteren. Hoofdstuk 2 verkent de impact van stotteren op kinderen in de voorschoolse leeftijd in termen van gezondheidsgerelateerde kwaliteit van leven (KVL). Ouders van 199 kinderen die deelnamen aan de RESTART-studie -een gerandomiseerde gecontroleerde studie waarin de behandeling van stotteren bij jonge kinderen is geëvalueerd- vulden verschillende vragenlijsten in over de KVL van hun kinderen. De gegevens die voorafgaand aan de behandeling werden verzameld, werden vergeleken met KVL gegevens van een groep niet stotterende Nederlandse kinderen (een zogenoemde vergelijkende populatie). De resultaten wijzen erop dat de KVL van kinderen die stotteren in de voorschoolse leeftijd vergelijkbaar is met die van kinderen die niet stotteren. Stotterernst en duur van het stotteren waren niet gerelateerd aan KVL scores. De conclusie van het hoofdstuk luidt dat over het algemeen de KVL van kinderen die stotteren niet aangedaan is. Een beperking van onze studie is dat slechts een klein aantal kinderen

met ernstig stotteren deelnamen. We adviseren daarom in vervolgonderzoek een groter cohort kinderen met ernstig stotteren op te nemen. Daarnaast is het relevant om te kijken naar het verloop van KVL in de tijd en om aanvullende parameters, zoals kenmerken van het kind of de omgeving, mee te nemen in vervolganalyses.

Hoofdstuk 3 laat zien dat, in geval van blijvend stotteren in de volwassen leeftijd, de impact op KVL groter wordt. Er namen 91 volwassenen deel aan de studie die in hoofdstuk 3 beschreven wordt. Zij werden voor een deel via ons eigen netwerk geworven en voor een deel via stottercentra en logopediepraktijken verspreid over Nederland. Door middel van een uitgebreide onderzoeksaanpak werd hun stotterernst, functioneren, KVL en copingstijl geëvalueerd. Ernstig stotteren bleek samen te hangen met lagere KVL scores. Een hogere stotterernst was niet alleen gerelateerd aan een lagere score op het spraakdomein van een van de KVL-instrumenten (HUI3), maar ook aan een lagere score op emotionele en sociale domeinen. Copingstijl hing even sterk samen met KVL als stotterernst. In het bijzonder was een hogere score op de emotie-gerichte subschaal van het CISS copinginstrument (wat een negatievere manier van omgaan met emoties weergeeft) gerelateerd aan een lagere KVL waarde.

In de studies beschreven in hoofdstuk 2 en 3 zijn generieke KVL vragenlijsten gebruikt. Dat wil zeggen dat ze KVL meten in termen die voor iedereen relevant zijn, dus ongeacht de aan- of afwezigheid van ziekten. Generieke KVL instrumenten kunnen daarom in diverse populaties gebruikt worden. Ziektespecifieke instrumenten daarentegen bevatten domeinen die relevant worden geacht voor mensen die een specifieke aandoening hebben. Het heeft de voorkeur om in KVL onderzoek zowel een generiek als ziektespecifiek instrument te gebruiken. Ten tijde van ons onderzoek was er echter geen stotterspecifiek KVL instrument voorhanden. Een aantal jaar daarvoor was de Engelse OASES voor volwassenen ontwikkeld en psychometrisch gevalideerd. Ondanks dat de OASES geen KVL instrument is, bevat het wel een onderdeel dat de invloed van stotteren op KVL meet. Het instrument geeft de ervaring met betrekking tot het stotteren weer vanuit het perspectief van de volwassene die stottert. Hoofdstuk 4 beschrijft het proces van het vertalen van de OASES naar het Nederlands en het onderzoek naar de psychometrische eigenschappen van het instrument. We concluderen dat de Nederlandse OASES voor volwassenen (OASES-A-D) een betrouwbaar en valide instrument is dat van waarde kan zijn in de klinische praktijk om de invloed van stotteren op het dagelijks leven te onderzoeken. Alle onderdelen van de OASES-A-D waren in staat om groepen deelnemers met een verschillend niveau van stotterernst van elkaar te onderscheiden. Verder onderzoek naar de test-hertestbetrouwbaarheid en mogelijke reductie van het aantal vragen wordt geadviseerd.

De hoofdstukken 5, 6 en 7 van dit proefschrift zijn gewijd aan de RESTART-studie (Rotterdam Evaluation study of Stuttering Therapy in preschool children- A Randomized Trial). In deze 18 maanden durende studie werden 199 kinderen in de leeftijd van 3 tot 6 jaar die minstens zes maanden stotterden op basis van loting toegewezen aan het Lidcombe Programma (LP) of behandeling gebaseerd op het Verwachtingen en Mogelijkheden model (RESTART-DCM behandeling). Behandeling volgens de DCM methode is sinds de jaren tachtig de Nederlandse standaard behandeling voor jonge kinderen die stotteren. RESTART-DCM behandeling is een *indirecte* benadering, wat wil zeggen dat behandeling gericht is op het werken aan kind- en omgevingsgerelateerde factoren waarvan verondersteld wordt dat ze de spraakvloeiendheid van het kind beïnvloeden. Het LP, de Australische standaardbehandeling, werd in 2000 geïntroduceerd in Nederland. Deze *directe* benadering maakt gebruik van operante gedragstherapeutische principes om de vloeiendheid van het spreken van het kind te beïnvloeden. In totaal deden 24 logopedisten mee aan de RESTART-studie, werkzaam in vrijgevestigde logopedische (stotter)centra verspreid over Nederland. Gezondheidsuitkomsten en kosten werden geëvalueerd voorafgaand en 3, 6, 12 en 18 maanden na aanvang van de behandeling.

Hoofdstuk 5 beschrijft de studie naar de vergelijkende effectiviteit van het LP en RESTART-DCM behandeling. De primaire uitkomstmaat was het percentage kinderen dat na 18 maanden niet meer stotterde. Secundaire uitkomstmaten, gemeten op de verschillende meetmomenten, waren de stotterfrequentie, het oordeel over de stotterernst gegeven door de ouders, therapeut en het kind zelf, KVL, emotionele en gedragsproblemen en de spreekattitude van het kind. Afgezien van een grotere verbetering van het spreken in de eerste drie maanden na behandeling met het LP waren de resultaten op de diverse uitkomstmaten voor beide behandelingen na 18 maanden gelijk. In de LP behandelgroep werden iets meer kinderen na 18 maanden geclassificeerd als niet stotterend in vergelijking met de RESTART-DCM groep (respectievelijk 76.5 en 71.4%), maar dit verschil was niet significant. We concluderen dat beide methoden effectief zijn in het behandelen van jonge kinderen die stotteren. Een opvallende uitkomst was dat in beide behandelgroepen de grootste verbetering in het spreken in de eerste drie maanden plaatsvond. De stotterfrequentie daalde hierna nog licht, terwijl op het meetmoment na 18 maanden ongeveer 30% van de kinderen nog in behandeling was.

Hoofdstuk 6 presenteert de resultaten van een economische evaluatie van het LP versus RESTART-DCM behandeling. De totale kosten in 18 maanden werden gerelateerd aan de volgende gezondheidsuitkomsten: 'Number needed to treat' (NNT: Het aantal kinderen dat gemiddeld genomen met het LP behandeld moet worden om één kind extra van het stotteren af te helpen, vergeleken met wanneer deze kinderen behandeld worden volgens de RESTART-DCM methode) en voor kwaliteit van leven gecorrigeerde levensjaren

(QALY's), gebaseerd op KVL scores. Omdat de KVL scores hoog waren aan het begin van de studie, werd slechts een kleine KVL winst behaald in de 18 maanden van de studie. Na 18 maanden waren de gezondheidsuitkomsten iets beter voor de LP groep, maar de verschillen waren statistisch en/of klinisch niet betekenisvol. De totale kosten waren iets hoger voor de LP groep. Aangezien er na 18 maanden een iets hoger percentage kinderen in de RESTART-DCM groep nog in behandeling was, zal het verschil in kosten op termijn mogelijk afnemen. Op basis van de NNT van 20 en het verschil in kosten van €168 worden de extra kosten per NNT geschat op €3360. Dit wil zeggen dat een extra investering van €3360 zal resulteren in één extra kind dat geclassificeerd word als 'niet stotterend' na behandeling met het LP in vergelijking tot RESTART-DCM behandeling. De extra kosten om extra verbetering in KVL te bewerkstelligen met het LP in vergelijking met de RESTART-DCM behandeling waren €10413 per extra V-QALY en €18617 per extra U-QALY. Deze ratio's impliceren een goede kosten-effectiviteit van het LP. Op basis van de uitkomsten die in hoofdstuk 5 en 6 gepresenteerd zijn, concluderen we dat het LP een goed alternatief is voor behandeling volgens de RESTART-DCM methode in de Nederlandse eerstelijnszorg.

Het is bekend dat beslissingen betreffende behandelkeuzes en implementatie van studieresultaten beïnvloed worden door ideeën, percepties en ervaringen van therapeuten. Daarom is een focus groep bijeenkomst gehouden met logopedisten die aan de RESTART-studie deelnamen. De resultaten hiervan zijn beschreven in hoofdstuk 7. De houdingen en overtuigingen van therapeuten ten opzichte van het LP en RESTART-DCM behandeling werden geëvalueerd, evenals hoe deze mogelijk veranderd zijn door deelname aan de RESTART-studie. In de eindfase van de studie bleken deelnemende therapeuten flexibeler en meer cliënt-gericht in hun keuzes voor een behandeling. Dit hing samen met de bevinding dat de mogelijkheden van beide behandelingen duidelijker waren geworden tijdens het onderzoek. Daar waar, voorafgaand aan het onderzoek, bijna alle logopedisten een voorkeur hadden voor een van beide behandelmethoden, was deze voorkeur in de eindfase van het onderzoek zo goed als verdwenen. Alle logopedisten gaven aan dat hun behandelkeuze vooral bepaald werd door kindgerelateerde factoren en door een eventuele voorkeur van ouders.

Hoofdstuk 8 bediscussieert drie thema's die voortkomen uit dit proefschrift en die betrekking hebben op (1) vroegtijdige interventie, (2) behandelkeuze en (3) verbetering van het therapieproces. Ten eerste onderschrijft de bevinding dat blijvend ernstig stotteren op de volwassen leeftijd samenhangt met een grotere impact op KVL de noodzaak van vroege interventie. Aangezien de KVL van kinderen in de voorschoolse leeftijd die stotteren nauwelijks verlaagd bleek, kan behandeling voorafgegaan worden door een periode waarin het stotteren van het kind gemonitord wordt, zoals ook in de huidige richtlijn voor stotteren wordt geadviseerd. Ten tweede is er op basis van onze resultaten geen

duidelijke voorkeur uit te spreken voor een van beide behandelingen (LP of RESTART-DCM behandeling). Een aanpak waarbij de logopedist samen met de ouders een keuze voor een behandeling maakt is aan te raden. Ten derde adviseren we het gebruik van de RESTART-DCM werkwijze door logopedisten die een kind volgens de DCM methode willen behandelen, te stimuleren en daarnaast de winst die behandeling na drie maanden met zich meebrengt in heroverweging te nemen. Hoofdstuk 8 bespreekt ook methodologische uitdagingen die samenhingen met de RESTART-studie en met het meten van de impact van stotteren op het dagelijks leven. In het bijzonder komt de validiteit van KVL instrumenten aan bod. Het hoofdstuk besluit met implicaties voor logopedisten en beleidsmakers en aanbevelingen voor verder onderzoek. Er is voldoende bewijs om de implementatie van het LP in de dagelijkse praktijk in Nederland aan te moedigen. Vanwege het kleine verschil in effecten en kosten tussen het LP en de RESTART-DCM behandeling, adviseren we om beide behandelingen in het verzekerde pakket te behouden. Toekomstig onderzoek moet uitwijzen of specifieke factoren bijdragen aan een positieve behandeluitkomst, of de behandelduur en/of –intensiteit verkort kan worden zonder de effectiviteit aan te tasten, en wat de lange termijn effecten zijn van behandeling in de voorschoolse leeftijd.

S

LIST OF ABBREVIATIONS

ANOVA	Analysis of Variance
AWS	Adults Who Stutter
AWNS	Adults Who do Not Stutter
CA scale	Clinical Assessment scale
CBCL	Child Behaviour Checklist
CEA	Cost-Effectiveness Analysis
CHQ-PF28	Child Health Questionnaire-Parent Form 28 items
CI	Confidence Interval
CISS	Coping Inventory for Stressful Situations
CISS-A	CISS Avoidance-oriented coping style
CISS-E	CISS Emotion-oriented coping style
CISS-T	CISS Task-oriented coping style
CUA	Cost-Utility Analysis
CWS	Children Who Stutter
DCM	Demands and Capacities Model
EBP	Evidence Based Practice
Erasmus MC	Erasmus Medical Center
ES	Effect Size
EQ-5D	EuroQol five-Dimensional
EQ-VAS	EuroQol Visual Analogue Scale
HRQOL	Health-Related Quality of Life
HUI3	Health Utility Index mark 3
ICER	Incremental Cost-Effectiveness Ratio
ICUR	Incremental Cost-Utility Ratio
ITQOL-97	Infant and Toddler Quality of Life Questionnaire 97 items
KiddyCAT	Communication Attitude Test- preschool child version
LP	Lidcombe Program for early stuttering intervention
NNT	Number Needed to Treat
NT group	adults who stutter who were Not in Therapy
OASES	Overall Assessment of the Speaker's Experience of Stuttering
OASES-A-D	Dutch version of the OASES for Adults
OR	Odds Ratio
QALY	Quality-Adjusted Life Year
QOL	Quality of Life
RCT	Randomized Controlled Trial
Resp	Respondent
RESTART	the Rotterdam Evaluation study of Stuttering Therapy in preschool children- A Randomized Trial

A

S24	Scale of of communication attitudes with 24 items
SA1	Sensitivity Analysis 1
SA-scale	Self-Assessment scale of speech
SD	Standard Deviation
SE	Standard Error
SF-36	medical outcomes study Short Form 36-Item health survey
SG	Standard Gamble
SLP	Speech-Language Pathologist
SPSS	Statistical Package of Social Sciences
SS	Syllables Stuttered
SSI-3	Stuttering Severity Instrument-3
T group	adults who stutter who were in Therapy
TSO	Time Since Onset
TTO	Time Trade-Off
U-QALYs	Quality-Adjusted Life Years based on measurement by the HUI3
V-QALYs	Quality-Adjusted Life Years based on measurement by the EQ-VAS
WTP	Willingness To Pay

PHD PORTFOLIO

Name PhD student: Caroline de Sonneville-Koedoot
EUR department: iBMG
Erasmus MC department: Ear, Nose and Throat clinic
PhD period: 2007-2015
Promotors: Prof. dr. Carin Uyl-de Groot
                         Prof. dr. Rob Baatenburg de Jong
Co-promotors: Dr. Elly Stolk
                         Dr. Marie-Christine Franken

PhD training

Health Technology Assessment, Radboud University (2007)
Academic Writing, Language & Training Centre, Erasmus University Rotterdam (2008)
Tutor training course Problem Based Learning (PGO), Institute of Psychology, Erasmus University Rotterdam (2008)
Course Basic Didactics, Risbo, Erasmus University Rotterdam (2008)
Regression analysis, NIHES winter school (2010)
iBMG didactic courses, De Loef Training, Erasmus University Rotterdam (2013)

Conferences

Podium presentations

IFA World Congress, Rio de Janeiro (2009)
Annual Congress NVLF, Ede-Wageningen (2009)
Congress on stuttering, Hogeschool Gent Vesalius, Gent (2009)
Annual Congress NVSST, Utrecht (2009)
ASHA Annual Convention, San Diego (2011)
Annual Congress NVSST, Utrecht (2012)
Annual Congress NVLF, Nieuwegein (2013)
10th Oxford Dysfluency Conference, Oxford (2014)

Poster presentations

NVLF80 Congress, Ede-Wageningen (2007)

Presentations at other meetings

Research Day ENT Erasmus MC, Rotterdam (2008, 2010)
Guest lectures Opleiding Logopedie HR, Rotterdam (2008-2015)
Staff meeting ENT Erasmus MC, Rotterdam (2009, 2013)

P

iMTA lunch lecture, iBMG, Rotterdam (2011)
Research meeting ENT Erasmus MC, Rotterdam (2011, 2012, 2014)
RESTART results meeting, Rotterdam (2012)
Symposium Stuttering KNAW, Amsterdam (2014)

Other conferences

European Symposium Fluency Disorders, Antwerp (2008)
ISPOR, Athens (2008)
LoLaHESG conference, Egmond aan Zee (2010) and Soesterberg (2011)

Teaching

Quality and efficiency in health care, tutor and lecturer (2008-2013)
General study skills, practicum and mentor of first-year student (2009-2014)
Supervisor and co-evaluator for bachelor and master theses iBMG and Opleiding Logopedie HR (2008-2012)
Writing and research skills for premaster students, practicum (2012-2013)

LIST OF SCIENTIFIC PUBLICATIONS

Publications in this thesis

Koedoot C, Bouwmans C, Franken MC, Stolk E. (2011). Quality of life in adults who stutter. J Commun Disord, 44, 429-443.

Koedoot C, Versteegh M, Yaruss JS. (2011). Psychometric evaluation of the Dutch translation of the Overall Assessment of the Speaker's Experience of Stuttering for adults (OASES-A-D). J Fluency Disord, 36(3), 222-230.

de Sonneville-Koedoot C, Stolk EA, Raat H, Bouwmans-Frijters C, Franken MC. (2014). Health-related quality of life of preschool children who stutter. J Fluency Disord, 42, 1-12.

de Sonneville-Koedoot C, Stolk E, Rietveld T, Franken MC. (2015). Direct versus Indirect Treatment for Preschool Children who Stutter: The RESTART Randomized Trial. PLoS ONE 10(7), e0133758.

de Sonneville-Koedoot C, Adams SA, Stolk EA, Franken MC. (2015). Perspectives of clinicians involved in the RESTART-study: Outcomes of a focus group. Am J Speech Lang Pathol. doi: 10.1044/2015_AJSLP-14-0215.

de Sonneville-Koedoot C, Bouwmans C, Franken MC, Stolk E. Economic evaluation of stuttering treatment in preschool children: the RESTART-study. Submitted.

Publications not in this thesis

Kanters TA, de Sonneville-Koedoot C, Redekop WK, Hakkaart L. (2013). Systematic review of available evidence on 11 high-priced inpatient orphan drugs. Orphanet J Rare Dis, 16(8), 124-1172-8-124.

Bouwmans C, de Sonneville C, Mulder CL, Hakkaart-van Roijen L. (2015). Employment and the associated impact on quality of life in people diagnosed with schizophrenia. Neuropsychiatr Dis Treat, 11, 2125-2142.

de Sonneville-Koedoot C, Hafkamp-de Groen E, van der Schroeff MPD, Stolk EA, Moll H, Jaddoe VWV, Hofman A, Raat H. The Health related Quality of Life of preschool children with AOM. Submitted.

P

ABOUT THE AUTHOR

Caroline de Sonneville-Koedoot was born in Rotterdam on the 15th of March 1979. She grew up in Oud-Beijerland. After her graduation from secondary school in 1997, she started studying Psychology at the Vrije Universiteit in Amsterdam, where she obtained a propaedeutic degree in 1998. From 1998 to 2002 she studied Speech and Language Therapy at the Hogeschool Rotterdam. Thereafter she started working as a speech-language therapist and started studying Speech and Language Pathology at the Radboud University in Nijmegen. She obtained her Master degree (cum laude) in 2006. In 2007 she joined the institute of Health Policy and Management at the Erasmus University (iBMG) in Rotterdam as a junior researcher, where she initiated a PhD thesis in collaboration with iBMG and the department of Ear, Nose and Throat of the Erasmus Medical Centre in Rotterdam. Caroline lives together with her husband Ben and their two children: Rosa (2012) and Sepp (2014).

A

DANKWOORD

Het is een heerlijk gevoel om na jaren dankwoorden gelezen te hebben nu eindelijk mijn eigen dankwoord te mogen schrijven. Ik wil heel graag iedereen bedanken die direct of indirect heeft bijgedragen aan de totstandkoming van dit proefschrift.

Allereerst mijn promotoren; oorspronkelijk Frans Rutten, daarna Carin Uyl-de Groot vanuit het iBMG, en Rob Baatenburg de Jong vanuit KNO. Bedankt voor de kans die ik kreeg om me te verdiepen in een onderwerp op het grensvlak van twee interessante vakgebieden. Carin, jij was altijd bereid mee te denken en hielp me de grote lijnen te blijven zien. Rob, in de afgelopen jaren hebben we de diverse uitkomstmaten en uitkomsten van studies bediscussieerd, meestal tijdens of na afloop van een van mijn presentaties. Jouw kritische vragen zetten me altijd weer aan het denken.

Elly en Marie-Christine, met jullie heb ik de bulk van het werk gedaan en alle hobbels die we onderweg tegenkwamen, genomen. Dank voor jullie begeleiding. MC, jouw ideeën liggen aan de basis van dit proefschrift. Ik bewonder jouw kennis van het vakgebied Stotteren en de bevlogenheid en gedrevenheid waarmee je je vak uitoefent. Dit heeft mij vaak gestimuleerd om nog een stapje verder te zetten. Elly, van jou heb ik veel over het vak gezondheidseconomie geleerd, zodat begrippen als HTA, QALYs en ICERs al snel heel gewoon voor me werden. Een rekensommetje laat zien dat wij samen in totaal ongeveer 3 jaar en 8 maanden zwanger zijn geweest gedurende mijn promotietraject. Met daarbij opgeteld ongeveer 15 maanden bevallingsverlof en nog eens tweeënhalf jaar verder ont-zwangeren vind ik dat we een knappe prestatie hebben geleverd. Jouw 'niet snel uit het veld geslagen' instelling en daarbij horende pragmatische aanpak brachten de rode draad weer helder naar voren wanneer ik dreigde te verzanden in details.

De leden van de promotiecommissie wil ik bedanken voor het lezen en beoordelen van mijn proefschrift en het opponeren bij de verdediging.

Bij het schrijven van alle artikelen voor dit proefschrift heb ik de prettige samenwerking met en de waardevolle input van anderen mogen ervaren. Toni, in Nijmegen heb jij me enthousiast gemaakt voor onderzoek binnen de spraak-taalpathologie. Ik ben blij dat je, ondanks dat je met emiraat bent, hebt mee willen werken in het lange traject dat tot een mooie publicatie van ons hoofdartikel heeft geleid. Scott, thank you very much for your contribution to Chapter 4. It was a genuine pleasure working with you. Matthijs, het artikel van hoofdstuk 4 was verreweg de snelste publicatie uit dit boekje. Jouw kennis en kunde hebben daar zeker aan bijgedragen! Sam, dank voor de prettige samenwerking en jouw waardevolle inbreng voor hoofdstuk 7; een geslaagd uitstapje naar het kwalitatieve onder-

D

zoek. Clazien, naar jouw gaat mijn speciale dank uit. Je bent coauteur van drie artikelen in dit proefschrift en hebt me met raad en daad bijgestaan tijdens het RESTART-project. Zo heb je bijvoorbeeld de inclusie van me overgenomen in de periode dat ik met verlof was. Veel dank daarvoor.

Het uitvoeren van een gerandomiseerd onderzoek is een hele klus. De medewerking van alle RESTART-logopedisten/stottertherapeuten was onmisbaar. Alies, Anne, Anneke, Brenda, Brunette, Caroline, Corrie, Durdana, Eeuwkje, Ellen, Esther, Fine, Hannie, Irma, Jeanette, Karin, Liesbeth, Lisette, Mark, Mary, Meina en Patricia: Jullie enthousiasme en inzet heeft ertoe geleid dat we, met een verlenging van anderhalf jaar, het toch voor elkaar hebben gekregen om bijna 200 kinderen te includeren. Ik kijk terug op vele momenten van prettig contact, bijvoorbeeld tijdens alle vrijdagmiddagen Intervisie, waarvoor jullie je praktijk moesten sluiten en naar Utrecht afreisden om onder het genot van wat lekkers uit een van de windstreken en met veel inzet (en Durdana als enthousiaste voorzitter) allerlei aan het onderzoek gerelateerde onderwerpen te bespreken. Met als een van laatste bijeenkomsten de focus groep meeting waar hoofdstuk 7 uit dit proefschrift op gebaseerd is. Als voorlopers in het veld zijn jullie een voorbeeld voor andere logopedisten. Chapeau voor jullie allen!

Judith, Maaike, Marjon, Robine, Sandra, Veronica, Wendy en Yoranique; dank jullie wel voor het scoren van de ruim 2700 audiosamples!!

Mijn bijzondere dank gaat uit naar alle kinderen en hun ouders die deelgenomen hebben aan de RESTART-studie. Dank voor jullie deelname aan het onderzoek, het maken van alle audio-opnamen thuis, het invullen van de vragenlijstboekjes en het soms speciaal naar de praktijk komen voor een meting.

Oud-collega's, ook al ben ik alweer even weg bij het iBMG, dankzij jullie was het altijd prettig om weer een dagje op de uni te zijn om aan m'n proefschrift te werken. Steef, verreweg de meeste jaren heb ik een kamer met jou gedeeld. Dank voor alle gesprekken tussen het werken door en je bereidheid om je werk te onderbreken voor weer eens een vraag over Excel, ftp-servers of wat voor computer gerelateerd onderwerp dan ook. Het 'vrijdag kroket-clubje' is inmiddels niet meer, maar gezellig dat we het contact in de vorm van iets meer culinaire etentjes weten te houden, Annemieke, Claudine, Liesbet, Renske, Saskia en Sophie!

Lieve vrienden, bedankt voor jullie interesse in mijn promotie en vooral voor alle gezelligheid en fijne gesprekken in de afgelopen jaren tijdens etentjes, (strand)wandelingen, dagjes speeltuin, avondjes Delft, en andere momenten die ik met jullie gedeeld heb. Car en Nens, super leuk om vandaag twee spraak-taalpathologen en vriendinnen naast me te hebben staan als paranimf!

Familie en schoonfamilie; dank voor jullie interesse en steun! Jenny, jouw kritische blik op m'n Engelse stukken hebben het niveau van een aantal hoofdstukken in dit proefschrift verhoogd, waarvoor veel dank! Pap en mam, jullie hebben me altijd gesteund. De laatste jaren met vele oppasdagen zodat ik ongestoord mijn proefschrift af kon ronden. Dank jullie wel voor alles!

Tot slot, het onmisbare thuisfront. Ben, het schrijven van dit proefschrift was soms een zware opgave. Dankjewel voor jouw altijd aanwezige steun en het kritisch maar geduldig meedenken in alle processen die hierbij kwamen kijken. Lieve Rosa en Sepp, mijn zonnetjes in huis, dank dat jullie er zijn!

D

REFERENCES

1. Sackett DL, Straus SE, Richardson WS, et al. (2000). Evidence-based medicine: How to practice and teach EBM (2nd ed). Edinburgh: Churchill Livingstone.

2. van Son N. (2003). Quickscan logopedieonderzoek: een verkenningsstudie naar de wensen en behoeften omtrent (toegepast) onderzoek en onderzoeksprogramma's op het vakgebied van de logopedie in Nederland. Stichting Logopedie Fonds.

3. Langevin M, Kully D. (2003). Evidence-based treatment of stuttering: III. Evidence-based practice in a clinical setting. J Fluency Disord, 28(3), 219-35.

4. Finn P. (2003). Evidence-based treatment of stuttering: II. Clinical significance of behavioral stuttering treatments. J Fluency Disord, 28(3), 209-17.

5. Bothe AK. (2003). Evidence-based treatment of stuttering: V. The art of clinical practice and the future of clinical research. J Fluency Disord, 28(3), 247-57.

6. Jones M, Onslow M, Packman A, Williams S, Ormond T, Schwarz I, et al. (2005). Randomised controlled trial of the Lidcombe programme of early stuttering intervention. BMJ, 331(7518), 659.

7. Bloodstein O, Bernstein-Ratner N. (2008). A handbook on stuttering. Boston, MA: Thompson Delmar Learning.

8. Guitar B. (2013). Stuttering: An integrated approach to its nature and treatment. Philadelphia, PA: Lippincott Williams & Wilkins.

9. Yaruss JS. (2010). Assessing quality of life in stuttering treatment outcomes research. J Fluency Disord, 35(3), 190-202.

10. Yairi E, Ambrose N. (2013). Epidemiology of stuttering: 21st century advances. J Fluency Disord, 38(2), 66-87.

11. Reilly S, Onslow M, Packman A, Cini E, Conway L, Ukoumunne OC, et al. (2013). Natural history of stuttering to 4 years of age: a prospective community-based study. Pediatrics, 132(3), 460-467.

12. Månsson H. (2000). Childhood stuttering: Incidence and development. J Fluency Disord, 25(1), 47-57.

13. Ambrose NG, Cox NJ, Yairi E. (1997). The genetic basis of persistence and recovery in stuttering. J Speech Lang Hear Res, 40(3), 567-580.

14. Kang C, Riazuddin S, Mundorff J, Krasnewich D, Friedman P, Mullikin JC, et al. (2010). Mutations in the lysosomal enzyme-targeting pathway and persistent stuttering. N Engl J Med, 362(8), 677-685.

15. Riaz N, Steinberg S, Ahmad J, Pluzhnikov A, Riazuddin S, Cox NJ, et al. (2005). Genomewide significant linkage to stuttering on chromosome 12. Am J Hum Genet, 76(4), 647-651.

16. Andrews G, Morris-Yates A, Howie P, Martin N. (1991). Genetic factors in stuttering confirmed. Arch Gen Psychiatry, 48(11): 1034.

17. Felsenfeld S, Kirk KM, Zhu G, Statham DJ, Neale MC, Martin NG. (2000). A study of the genetic and environmental etiology of stuttering in a selected twin sample. Behav Genet, 30(5), 359-366.

18. Felsenfeld S, Plomin R. (1997). Epidemiological and offspring analyses of developmental speech disorders using data from the Colorado Adoption Project. J Speech Lang Hear Res, 40(4), 778-791.

19. Domingues CE, Olivera CM, Oliveira BV, Juste FS, Andrade CR, Giacheti CM, et al. (2014). A genetic linkage study in Brazil identifies a new locus for persistent developmental stuttering on chromosome 10. Genet Mol Res, 13(1), 2094-2101.

R

20. Han TU, Park J, Domingues CF, Moretti-Ferreira D, Paris E, Sainz E, et al. (2014). A study of the role of the FOXP2 and CNTNAP2 genes in persistent developmental stuttering. Neurobiol Dis, 6923-31.

21. Raza MH, Gertz EM, Mundorff J, Lukong J, Kuster J, Schaffer AA, et al. (2013). Linkage analysis of a large African family segregating stuttering suggests polygenic inheritance. Hum Genet, 132(4), 385-396.

22. Raza MH, Amjad R, Riazuddin S, Drayna D. (2012). Studies in a consanguineous family reveal a novel locus for stuttering on chromosome 16q. Hum Genet, 131(2), 311-313.

23. Beal DS, Lerch JP, Cameron B, Henderson R, Gracco VL, De Nil LF. (2015). The trajectory of gray matter development in Broca's area is abnormal in people who stutter. Front Hum Neurosci, 989.

24. Chang SE, Zhu DC, Choo AL, Angstadt M. (2015). White matter neuroanatomical differences in young children who stutter. Brain, 138(3), 694-711.

25. Beal DS, Quraan MA, Cheyne DO, Taylor MJ, Gracco VL, De Nil LF. (2011). Speech-induced suppression of evoked auditory fields in children who stutter. Neuroimage, 54(4), 2994-3003.

26. Jansson-Verkasalo E, Eggers K, Järvenpää A, Suominen K, Van den Bergh B, De Nil L, et al. (2014). Atypical central auditory speech-sound discrimination in children who stutter as indexed by the mismatch negativity. J Fluency Disord, 411-11.

27. Chang SE, Zhu DC. (2013). Neural network connectivity differences in children who stutter. Brain, 136(12): 3709-3726.

28. Byrd CT, Conture EG, Ohde RN. (2007). Phonological priming in young children who stutter: holistic versus incremental processing. Am J Speech Lang Pathol, 16(1), 43-53.

29. Zackheim CT, Conture EG. (2003). Childhood stuttering and speech disfluencies in relation to children's mean length of utterance: a preliminary study. J Fluency Disord, 28(2), 115-41.

30. Yaruss JS. (1999). Utterance length, syntactic complexity, and childhood stuttering. J Speech Lang Hear Res, 42(2), 329-344.

31. Anderson JD, Conture EG. (2000). Language Abilities of Children Who Stutter: a Preliminary Study. J Fluency Disord, 25(4), 283-304.

32. Anderson JD, Pellowski MW, Conture EG. (2005). Childhood stuttering and dissociations across linguistic domains. J Fluency Disord, 30(3), 219-253.

33. Sasisekaran J. (2014). Exploring the link between stuttering and phonology: a review and implications for treatment. Semin Speech Lang, 35(2), 95-113.

34. Yairi E, Ambrose NG. (2005). Early Childhood Stuttering: For Clinicians by Clinicians. Austin, TX: Pro-Ed.

35. Andrews G, Harris M. (1964). The syndrome of stuttering. London: Heinemann.

36. Craig A, Hancock K, Tran Y, Craig M, Peters K. (2002). Epidemiology of stuttering in the community across the entire life span. J Speech Lang Hear Res, 45(6), 1097-1105.

37. Dworzynski K, Remington A, Rijsdijk F, Howell P, Plomin R. (2007). Genetic etiology in cases of recovered and persistent stuttering in an unselected, longitudinal sample of young twins. Am J Speech Lang Pathol, 16(2), 169-178.

38. Yairi E, Ambrose NG, Paden EP, Throneburg RN. (1996). Predictive factors of persistence and recovery: pathways of childhood stuttering. J Comm Disord, 29(1), 51-77.

39. Blood GW, Blood IM. (2007). Preliminary study of self-reported experience of physical aggression and bullying of boys who stutter: relation to increased anxiety. Percept Mot Skills, 104(3 Pt 2), 1060-1066.

40. Iverach L, Rapee RM. (2014). Social anxiety disorder and stuttering: current status and future directions. J Fluency Disord, 4069-82.

41. Craig A, Blumgart E, Tran Y. (2009). The impact of stuttering on the quality of life in adults who stutter. J Fluency Disord, 34(2), 61-71.

42. Klein JF, Hood SB. (2004). The impact of stuttering on employment opportunities and job performance. Journal of fluency disorders, 29(4), 255-73.

43. Bricker-Katz G, Lincoln M, Cumming S. (2013). Stuttering and work life: an interpretative phenomenological analysis. J Fluency Disord, 38(4), 342-355.

44. Ingham RJ, Cordes AK. (1998). Treatment decisions for young children who stutter: Further concerns and complexities. Am J Speech Lang Pathol, 7(3), 10.

45. Blomgren M. (2013). Behavioral treatments for children and adults who stutter: a review. Psychol Res Behav Manag, 69-19.

46. Ward D. (2013). Risk factors and stuttering: evaluating the evidence for clinicians. J Fluency Disord, 38(2), 134-140.

47. Starkweather CW, Gottwald SR. (1990). The demands and capacities model II: Clinical applications. J Fluency Disord, 15(3), 143-157.

48. Onslow M, Packman A, Harrison E. (2003). The Lidcombe Program of Early Stuttering Intervention: A Clinicians's Guide. Austin, Texas: Pro-ed.

49. Lincoln MA, Onslow M. (1997). Long-term outcome of early intervention for stuttering. Am J Speech Lang Pathol, 6(1), 51.

50. Onslow M, Costa L, Rue S. (1990). Direct early intervention with stuttering: Some preliminary data. J Speech Hear Disord, 55(3), 405.

51. Onslow M, Andrews C, Lincoln M. (1994). A control/experimental trial of an operant treatment for early stuttering. J Speech Hear Res, 37(6), 1244-1259.

52. Gregory HH, Campbell JH, Gregory CB, Hill DG. (2003). Stuttering therapy: Rationale and procedures. Boston: Allyn & Bacon.

53. Smith A, Kelly E. (1997). Stuttering: A dynamic, multifactorial model. In: RF Curlee & GM Siegel (Eds.). Nature and treatment of stuttering: New directions(2nd ed). p. 2204-217.   Boston: Allyn & Bacon.

54. Starkweather C. (1997). Therapy for younger children. In: RF Curlee & GM Siegel (Eds.). Nature and treatment of stuttering: New directions(2nd ed). p. 257-279.   Boston: Allyn & Bacon.

55. Starkweather CW. (2002). The epigenesis of stuttering. J Fluency Disord, 27(4), 269-87.

56. Franken MC, Kielstra-Van der Schalk CJ, Boelens H. (2005). Experimental treatment of early stuttering: a preliminary study. J Fluency Disord, 30(3), 189-199.

57. Rustin L, Botterill W, Kelman E. (1996). Assessment and therapy for young dysfluent children: Family interaction. San Diego: Singular Publishing Group.

58. Kelman E, Nicholas A. (2008). Practical intervention for early childhood stammering: Palin PCI approach. Speechmark.

59. Yaruss JS, Coleman C, Hammer D. (2006). Treating preschool children who stutter: description and preliminary evaluation of a family-focused treatment approach. Lang Speech Hear Serv Sch, 37(2), 118-136.

60. Millard SK, Nicholas A, Cook FM. (2008). Is parent-child interaction therapy effective in reducing stuttering? J Speech Lang Hear Res, 51(3), 636-650.

R

61. Millard SK, Edwards S, Cook FM. (2009). Parent-child interaction therapy: Adding to the evidence. Int J Speech Lang Pathol, 11(1), 61-76.

62. Meyers SC, Woodford LL, Callisto TA. (1992). The Fluency Development System for Young Children (ages 2-9 Years). Callisto, TA (Ed.). United Educational Services.

63. Riley GD, Ingham JC. (2000). Acoustic duration changes associated with two types of treatment for children who stutter. J Speech Lang Hear Res, 43(4), 965-978.

64. Ryan BP. (2001). Programmed therapy for stuttering in children and adults. Springfield, IL: Charles C Thomas Pub Ltd.

65. Pertijs M. [Internet]. (2014). Evidence-based richtlijn ontwikkelingsstotteren bij kinderen en volwassenen. Kwaliteit van bewijs voor het effect van behandeling van jonge stotterende kinderen. Available at: http://www.nedverstottertherapie.nl.

66. Nye C, Vanryckeghem M, Schwartz JB, Herder C, Turner HM,3rd, Howard C. (2013). Behavioral stuttering interventions for children and adolescents: a systematic review and meta-analysis. J Speech Lang Hear Res, 56(3), 921-932.

67. Harris V, Onslow M, Packman A, Harrison E, Menzies R. (2002). An experimental investigation of the impact of the Lidcombe Program on early stuttering. J Fluency Disord, 27(3), 203-13.

68. Harrison E, Onslow M, Menzies R. (2004). Dismantling the Lidcombe Program of early stuttering intervention: verbal contingencies for stuttering and clinical measurement. Int J Lang Commun Disord, 39(2), 257-267.

69. Lattermann C, Euler HA, Neumann K. (2008). A randomized control trial to investigate the impact of the Lidcombe Program on early stuttering in German-speaking preschoolers. J Fluency Disord, 33(1), 52-65.

70. Lewis C, Packman A, Onslow M, Simpson JM, Jones M. (2008). A phase II trial of telehealth delivery of the Lidcombe Program of Early Stuttering Intervention. Am J Speech Lang Pathol, 17(2), 139-149.

71. Brutten G, Shoemaker D. (1974). Speech situation checklist. Carbondale, IL: Speech Clinic, Southern Illinois University.

72. Andrews G, Cutler J. (1974). Stuttering therapy: the relation between changes in symptom level and attitudes. J Speech Hear Disord, 39(3), 312-319.

73. Erickson RL. (1969). Assessing communication attitudes among stutterers. J Speech Hear Res, 12(4), 711-724.

74. Riley J, Riley G, Maguire G. (2004). Subjective Screening of Stuttering severity, locus of control and avoidance: research edition. J Fluency Disord, 29(1), 51-62.

75. Franic DM, Bothe AK. (2008). Psychometric evaluation of condition-specific instruments used to assess health-related quality of life, attitudes, and related constructs in stuttering. Am J Speech Lang Pathol, 17(1), 60-80.

76. Yaruss JS, Quesal RW. (2004). Partnerships between clinicians, researchers, and people who stutter in the evaluation of stuttering treatment outcomes. Stammering Research, 11-15.

77. Craig A. (2010). The association between quality of life and stuttering. J Fluency Disord, 35(3), 159-160.

78. Ingham JC. (2003). Evidence-based treatment of stuttering: I. Definition and application. J Fluency Disord, 28(3), 197-206.

79. Wilson IB, Cleary PD. (1995). Linking clinical variables with health-related quality of life. A conceptual model of patient outcomes. JAMA, 273(1), 59-65.

80. World Health Organization. (2001). International classification of functioning, disability and health (ICF). Geneva: World Health Organization.

81. Roddey TS, Cook KF, O'Malley KJ, Gartsman GM. (2005). The relationship among strength and mobility measures and self-report outcome scores in persons after rotator cuff repair surgery: impairment measures are not enough. J Shoulder Elbow Surg, 14(1 Suppl S), 95S-98S.

82. Hunt SM, McKenna SP. (1993). Measuring quality of life in psychiatry. Quality of life assessment: Key issues in the 1990s: Springer Netherlands, p. 343-354.

83. Crichton-Smith I. (2002). Communicating in the real world: accounts from people who stammer. J Fluency Disord, 27(4), 333-352.

84. Hayhow R, Cray AM, Enderby P. (2002). Stammering and therapy views of people who stammer. J Fluency Disord, 27(1), 1-16.

85. Klompas M, Ross E. (2004). Life experiences of people who stutter, and the perceived impact of stuttering on quality of life: personal accounts of South African individuals. J Fluency Disord, 29(4), 275-305.

86. Yaruss JS, Quesal RW. (2006). Overall Assessment of the Speaker's Experience of Stuttering (OASES): documenting multiple outcomes in stuttering treatment. J Fluency Disord, 31(2), 90-115.

87. Feeney R, Desha L, Ziviani J, Nicholson JM. (2012). Health-related quality-of-life of children with speech and language difficulties: A review of the literature. Int J Speech Lang Pathol, 14(1): 59-72.

88. Drummond MF, Sculpher MJ, Torrance GW, O'Brien BJ, Stoddart GL. (2005). Methods for the economic evaluation of health care programmes. Oxford: Oxford University Press.

89. Raat H, Mohangoo AD, Grootenhuis MA. (2006). Pediatric health-related quality of life questionnaires in clinical trials. Curr Opin Allergy Clin Immunol, 6(3), 180-185.

90. Solans M, Pane S, Estrada MD, Serra-Sutton V, Berra S, Herdman M, et al. (2008). Health-related quality of life measurement in children and adolescents: a systematic review of generic and disease-specific instruments. Value Health, 11(4), 742-764.

91. Koedoot C, Bouwmans C, Franken MC, Stolk E. (2011). Quality of life in adults who stutter. J Commun Disord, 44(4): 429-443.

92. Gooding S, Davis S. (2004). Indications by parents of 'environmental' factors that differ between children who stutter and controls. Stammering Research, 121-23.

93. Chun RY, Mendes CD, Yaruss JS, Quesal RW. (2010). The impact of stuttering on quality of life of children and adolescents. Pro Fono, 22(4), 567-570.

94. Cook S, Donlan C, Howell P. (2013). Stuttering severity, psychosocial impact and lexical diversity as predictors of outcome for treatment of stuttering. J Fluency Disord, 38(2), 124-133.

95. Kawai N, Healey EC, Nagasawa T, Vanryckeghem M. (2012). Communication attitudes of Japanese school-age children who stutter. J Commun Disord, 45(5), 348-354.

96. Vanryckeghem M, Brutten GJ, Hernandez LM. (2005). A comparative investigation of the speech-associated attitude of preschool and kindergarten children who do and do not stutter. J Fluency Disord, 30(4), 307-318.

97. Smith KA, Iverach L, O'Brian S, Kefalianos E, Reilly S. (2014). Anxiety of children and adolescents who stutter: a review. J Fluency Disord, 4022-34.

98. van der Merwe B, Robb MP, Lewis JG, Ormond T. (2011). Anxiety measures and salivary cortisol responses in preschool children who stutter. Contemp Issues Commun Sci Disord., 38, 1-10.

99. Boey RA, Van de Heyning PH, Wuyts FL, Heylen L, Stoop R, De Bodt MS. (2009). Awareness and reactions of young stuttering children aged 2-7 years old towards their speech disfluency. J Commun Disord, 42(5), 334-346.

R

100. Conture EG, Kelly EM, Walden TA. (2013). Temperament, speech and language: An overview. J Commun Disord, 46(2), 125-142.

101. Kefalianos E, Onslow M, Block S, Menzies R, Reilly S. (2012). Early stuttering, temperament and anxiety: two hypotheses. J Fluency Disord, 37(3), 151-163.

102. Kefalianos E, Onslow M, Ukoumunne O, Block S, Reilly S. (2014). Stuttering, temperament, and anxiety: data from a community cohort ages 2-4 years. J Speech Lang Hear Res, 57(4), 1314-1322.

103. Ambrose NG, Yairi E. (1994). The development of awareness of stuttering in preschool children. J Fluency Disord, 19(4), 229-245.

104. Ezrati-Vinacour R, Platzky R, Yairi E. (2001). The young child's awareness of stuttering-like disfluency. J Speech Lang Hear Res, 44(2), 368-380.

105. Langevin M, Packman A, Onslow M. (2009). Peer responses to stuttering in the preschool setting. Am J Speech Lang Pathol, 18(3), 264-276.

106. Langevin M, Packman A, Onslow M. (2010). Parent perceptions of the impact of stuttering on their preschoolers and themselves. J Commun Disord, 43(5), 407-423.

107. Raat H, Landgraf JM, Oostenbrink R, Moll HA, Essink-Bot ML. (2007). Reliability and validity of the Infant and Toddler Quality of Life Questionnaire (ITQOL) in a general population and respiratory disease sample. Qual Life Res, 16(3), 445-460.

108. Raat H, Botterweck AM, Landgraf JM, Hoogeveen WC, Essink-Bot ML. (2005). Reliability and validity of the short form of the child health questionnaire for parents (CHQ-PF28) in large random school based and general population samples. J Epidemiol Community Health, 59(1), 75-82.

109. HealthActCHQ. (2007). Confidential scoring rules Infant Toddler Quality of Life Questionnaire-97 (ITQOL-97). Boston, MA: HealthActCHQ.

110. Landgraf JM, Abetz L, Ware JE. (1996). Child Health Questionnaire (CHQ): A user's manual. Health Institute.

111. Feeny D, Furlong W, Torrance GW, Goldsmith CH, Zhu Z, DePauw S, et al. (2002). Multiattribute and single-attribute utility functions for the health utilities index mark 3 system. Med Care, 40(2), 113-128.

112. Horsman J, Furlong W, Feeny D, Torrance G. (2003). The Health Utilities Index (HUI): concepts, measurement properties and applications. Health Qual Life Outcomes, 154.

113. EuroQol Group. EQ-5D User Guide. [Internet]. (2009). Basic information on how to use EQ-5D.

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

115. Brunner HI, Klein-Gitelman MS, Miller MJ, Trombley M, Baldwin N, Kress A, et al. (2004). Health of children with chronic arthritis: relationship of different measures and the quality of parent proxy reporting. Arthritis Rheum, 51(5), 763-773.

116. Riley GD ed. (1994). Stuttering severity instrument for children and adults. Austin: Pro-Ed.

117. Cohen J. (1988). Statistical power analysis for the behavioral sciences. 2nd ed. Hillsdale, NJ: Lawrence Erlbaum.

118. Varni JW, Limbers CA, Burwinkle TM. (2007). Parent proxy-report of their children's health-related quality of life: an analysis of 13,878 parents' reliability and validity across age subgroups using the PedsQL 4.0 Generic Core Scales. Health Qual Life Outcomes, 52.

119. Bannink N, Maliepaard M, Raat H, Joosten KF, Mathijssen IM. (2010). Health-related quality of life in children and adolescents with syndromic craniosynostosis. J Plast Reconstr Aesthet Surg, 63(12), 1972-1981.

120. Spuijbroek AT, Oostenbrink R, Landgraf JM, Rietveld E, de Goede-Bolder A, van Beeck EF, et al. (2011). Health-related quality of life in preschool children in five health conditions. Qual Life Res, 20(5), 779-786.

121. Addington-Hall J, Kalra L. (2001). Who should measure quality of life? BMJ, 322(7299), 1417-1420.

122. Streisand R, Braniecki S, Tercyak KP, Kazak AE. (2001). Childhood illness-related parenting stress: the pediatric inventory for parents. J Pediatr Psychol, 26(3), 155-162.

123. Theunissen NC, Vogels TG, Koopman HM, Verrips GH, Zwinderman KA, Verloove-Vanhorick SP, et al. (1998). The proxy problem: child report versus parent report in health-related quality of life research. Qual Life Res, 7(5), 387-397.

124. Rebok G, Riley A, Forrest C, Starfield B, Green B, Robertson J, et al. (2001). Elementary school-aged children's reports of their health: a cognitive interviewing study. Qual Life Res, 10(1), 59-70.

125. Vetter TR, Bridgewater CL, McGwin G,Jr. (2012). An observational study of patient versus parental perceptions of health-related quality of life in children and adolescents with a chronic pain condition: who should the clinician believe? Health Qual Life Outcomes, 1085-7525-10-85.

126. Davis S, Howell P, Cooke F. (2002). Sociodynamic relationships between children who stutter and their non-stuttering classmates. J Child Psychol Psychiatry, 43(7), 939-947.

127. Messenger M, Onslow M, Packman A, Menzies R. (2004). Social anxiety in stuttering: measuring negative social expectancies. J Fluency Disord, 29(3), 201-12.

128. Kraaimaat FW, Vanryckeghem M, Van Dam-Baggen R. (2002). Stuttering and social anxiety. Journal of fluency disorders, 27(4), 319-30.

129. Schneier FR, Wexler KB, Liebowitz MR. (1997). Social phobia and stuttering. Am J Psychiatry, 154(1), 131.

130. Menzies RG, Onslow M, Packman A, O'Brian S. (2009). Cognitive behavior therapy for adults who stutter: a tutorial for speech-language pathologists. J Fluency Disord, 34(3), 187-200.

131. Cruice M, Worrall L, Hickson L. (2010). Health-related quality of life in people with aphasia: implications for fluency disorders quality of life research. J Fluency Disord, 35(3), 173-189.

132. Andrade CR, Sassi FC, Juste FS, Ercolin B. (2008). Quality of life of individuals with persistent developmental stuttering. Pro Fono, 20(4), 219-224.

133. Brazier J, Ratcliffe J, Tsuchiya A, Salomon J. (2007). Measuring and valuing health benefits for economic evaluation. Oxford: Oxford University Press.

134. Torrance GW, Furlong W, Feeny D. (2002). Health utility estimation. Expert Rev Pharmacoecon Outcomes Res, 2(2), 99-108.

135. Coons SJ, Rao S, Keininger DL, Hays RD. (2000). A comparative review of generic quality-of-life instruments. Pharmacoeconomics, 17(1), 13-35.

136. Brazier J, Roberts J, Deverill M. (2002). The estimation of a preference-based measure of health from the SF-36. J Health Econ, 21(2), 271-292.

137. Bramlett RE, Bothe AK, Franic DM. (2006). Using preference-based measures to assess quality of life in stuttering. J Speech Lang Hear Res, 49(2), 381-94.

138. Busschbach JJ, Rikken B, Grobbee DE, De Charro FT, Wit JM. (1998). Quality of life in short adults. Horm Res, 49(1), 32-38.

R

139. Guyatt GH, Naylor CD, Juniper E, Heyland DK, Jaeschke R, Cook DJ. (1997). Users' guides to the medical literature. XII. How to use articles about health-related quality of life. Evidence-Based Medicine Working Group. JAMA, 277(15), 1232-1237.

140. Guyatt GH, Veldhuyzen Van Zanten SJ, Feeny DH, Patrick DL. (1989). Measuring quality of life in clinical trials: a taxonomy and review. CMAJ, 140(12), 1441-1448.

141. O'Brian S, Packman A, Onslow M. (2004). Self-rating of stuttering severity as a clinical tool. Am J Speech Lang Pathol, 13(3), 219-26.

142. Huinck W, Rietveld T. (2007). The validity of a simple outcome measure to assess stuttering therapy. Folia Phoniatr Logop, 59(2), 91-99.

143. Plexico LW, Manning WH, Levitt H. (2009). Coping responses by adults who stutter: part I. Protecting the self and others. J Fluency Disord, 34(2), 87-107.

144. Plexico L, Manning WH, Levitt H. (2009). Coping responses by adults who stutter: part II. Approaching the problem and achieving agency. J Fluency Disord, 34(2), 108-126.

145. Vanryckeghem M, Brutten GJ, Uddin N, Van Borsel J. (2004). A comparative investigation of the speech-associated coping responses reported by adults who do and do not stutter. J Fluency Disord, 29(3), 237-50.

146. Parker JDA, Endler NS. (1992). Coping with coping assessment: A critical review. Eur J Personality, 6321-321.

147. de Ridder D, Geenen R, Kuijer R, van Middendorp H. (2008). Psychological adjustment to chronic disease. Lancet, 372(9634), 246-255.

148. Dolan, P. (2008). Stuttering into happiness. Oxford Disfluency Conference.

149. de Ridder D, van Heck GL. (2004). Coping inventory for stressful situations. CISS Handleiding.

150. Dolan P. (1997). Modeling valuations for the EuroQol health states. Med Care, 351095-1108.

151. Oostenbrink R, A Moll HA, Essink-Bot ML. (2002). The EQ-5D and the Health Utilities Index for permanent sequelae after meningitis: a head-to-head comparison. J Clin Epidemiol, 55(8), 791-799.

152. Lazarus RS. (1993). Coping theory and research: Past, present, and future. Psychosom Med, 55, 234-234.

153. Doctor JN, Bleichrodt H, Lin HJ. (2008). Health Utility Bias: A Systematic Review and Meta-Analytic Evaluation. Med Decis Making, 30, 58-67.

154. Kopec JA, Willison KD. (2003). A comparative review of four preference-weighted measures of health-related quality of life. J Clin Epidemiol, 56(4), 317-325.

155. Lamers LM, Bouwmans CA, van Straten A, Donker MC, Hakkaart L. (2006). Comparison of EQ-5D and SF-6D utilities in mental health patients. Health Econ, 15(11), 1229-1236.

156. McCabe C, Brazier J, Gilks P, Tsuchiya A, Roberts J, O'Hagan A, et al. (2006). Using rank data to estimate health state utility models. J Health Econ, 25(3), 418-431.

157. Macdonald G, Leary MR. (2005). Why does social exclusion hurt? The relationship between social and physical pain. Psychol Bull, 131(2), 202-223.

158. Lazarus RS, Folkman S. (1984). Stress, appraisal, and coping. New York, NY: Springer Publishing Company.

159. Yaruss JS, Quesal RW. (2008). OASES: Overall Assessment of the Speaker's Experience of Stuttering. Bloomington, MN: Pearson Assessments.

160. Brutten G, Vanryckeghem M. (2003). Behavior Assessment Battery: A multi-dimensional and evidence-based approach to diagnostic and therapeutic decision making for adults who stutter. S-24 modification of Erickson's scale of communication attitudes revised 2003. Destelbergen, Belgium: Stichting Integratie Gehandicapten & Acco Publishers.

161. Cummins RA. (2010). Fluency disorders and life quality: subjective wellbeing vs. health-related quality of life. J Fluency Disord, 35(3), 161-172.

162. Yaruss J, Quesal R, Coleman C. (2010). Overall assessment of the speaker's experience of stuttering: Ages 13–17 (OASES-T) Response form. Bloomington, MN: Pearson Assessments.

163. Yaruss JS, Coleman G, Quesal RW. (2010). Overall Assessment of the Speakers Experience of Stuttering: Ages 7-12 (OASES-S) Response form. Bloomington, MN: Pearson Assessments.

164. Bricker-Katz G, Lincoln M, McCabe P. (2009). A life-time of stuttering: how emotional reactions to stuttering impact activities and participation in older people. Disabil Rehabil, 31(21), 1742-1752.

165. Mulcahy K, Hennessey N, Beilby J, Byrnes M. (2008). Social anxiety and the severity and typography of stuttering in adolescents. J Fluency Disord, 33(4), 306-319.

166. Cream A, O'Brian S, Jones M, Block S, Harrison E, Lincoln M, et al. (2010). Randomized controlled trial of video self-modeling following speech restructuring treatment for stuttering. J Speech Lang Hear Res, 53(4): 887-897.

167. Metten C, Zückner H, Rosenberger S. (2007). Evaluation of an intensive stuttering treatment for children and adolescents. Sprache, Stimme, Gehör, 31(2), 1-10.

168. Yaruss JS, Quesal RW. (2009, Nov). OASES-A: Translation & validation for multiple languages & cultures. Poster presented at the Annual Convention of the American Speech-Language-Hearing Association, New Orleans, LA.

169. Yaruss JS. (2001). Evaluating treatment outcomes for adults who stutter. J Commun Disord, 34(1-2), 163-182.

170. Herdman M, Fox-Rushby J, Rabin R, Badia X, Selai C. (2003). Producing other language versions of the EQ-5D. In: The measurement and valuation of health status using EQ-5D: A European perspective. Dodrecht: Kluwer Academic Publishers.

171. Beaton DE, Bombardier C, Guillemin F, Ferraz MB. (2000). Guidelines for the process of cross-cultural adaptation of self-report measures. Spine (Phila Pa 1976), 25(24), 3186-3191.

172. Huinck W, Rietveld T. (2007). The validity of a simple outcome measure to assess stuttering therapy. Folia Phoniatr Logop, 59(2), 91-99.

173. Peterson RA. (1994). A meta-analysis of Cronbach's coefficient alpha. Journal of consumer research, 381-391.

174. Nunnally JC. (1978). Psychometric theory (2nd ed.). New York, NY: McGraw-Hill.

175. Streiner DL, Norman GR, Cairney J. (2003). Health measurement scales: a practical guide to their development and use (3rd ed.). Oxford: Oxford University Press.

176. Hinkle DE, Wiersma W, Jurs SG. (1998). Applied statistics for the behavioral sciences. (4th ed). Boston, MA: Houghton Mifflin Company.

177. Kim, W, Yaruss JS. (2008, Nov). The relationship between the impact of stuttering and age. Poster presented at the Annual Convention of the American Speech-Language-Hearing Association, Chicago, IL.

178. Pallant JF, Tennant A. (2007). An introduction to the Rasch measurement model: an example using the Hospital Anxiety and Depression Scale (HADS). Br J Clin Psychol, 46(Pt 1), 1-18.

R

179. Tennant A, McKenna SP, Hagell P. (2004). Application of Rasch analysis in the development and application of quality of life instruments. Value Health, 7 Suppl 1S22-6.

180. Sommer M, Koch MA, Paulus W, Weiller C, Buchel C. (2002). Disconnection of speech-relevant brain areas in persistent developmental stuttering. Lancet, 360(9330), 380-383.

181. Chang SE, Erickson KI, Ambrose NG, Hasegawa-Johnson MA, Ludlow CL. (2008). Brain anatomy differences in childhood stuttering. Neuroimage, 39(3), 1333-1344.

182. Rautakoski P, Hannus T, Simberg S, Sandnabba NK, Santtila P. (2012). Genetic and environmental effects on stuttering: a twin study from Finland. J Fluency Disord, 37(3), 202-210.

183. Guitar B, Conture EG. [Internet]. (2013). The child who stutters: to the pediatrician. 5th edition. Publication no. 0023. Memphis, TN: Stuttering Foundation of America. Available: http://www.stutteringhelp.org.

184. O'Brian S, Onslow M. (2011). Clinical management of stuttering in children and adults. BMJ, 342d3742.

185. Onslow M, Menzies RG, Packman A. (2001). An operant intervention for early stuttering. The development of the Lidcombe program. Behav Modif, 25(1), 116-139.

186. Jones M, Onslow M, Packman A, O'Brian S, Hearne A, Williams S, et al. (2008). Extended follow-up of a randomized controlled trial of the Lidcombe Program of Early Stuttering Intervention. Int J Lang Commun Disord, 43(6), 649-661.

187. Gore K. The Australian Budget Lidcombe Proposal Debate: A Primer in Memes. [Internet]. (2015). Available at: http://www.speechirl.com/blog/the-australian-budget-lidcombe-proposal-debate-a-primer-in-memes.

188. Franken MC, Putker-de Bruijn D. [Internet]. (2014). RESTART-DCM Method. Treatment protocol developed within the scope of the ZonMW project Cost-effectiveness of the Demands and Capacities Model based treatment compared to the Lidcombe programme of early stuttering intervention: Randomised trial. Available at: http://www.nedverstottertherapie.nl/.

189. Evans S, Royston P, Day S. [Internet]. (2004). Minim: allocation by minimisation in clinical trials. Available at: http://www-users.york.ac.uk/~ mb55/guide/minim.htm.

190. O'Callaghan CA. (2014). OxMaR: open source free software for online minimization and randomization for clinical trials. PLoS One, 9(10), e110761.

191. Clark CE, Conture EG, Walden TA, Lambert WE. (2013). Speech sound articulation abilities of preschool-age children who stutter. J Fluency Disord, 38(4), 325-341.

192. Ingham, R. J. & Riley,G. (1998). Guidelines for Documentation of Treatment Efficacy for Young Children who stutter. J Speech Lang Hear R, 41753-770.

193. Sawyer J, Yairi E. (2006). The effect of sample size on the assessment of stuttering severity. Am J Speech Lang Pathol, 15(1), 36-44.

194. Achenbach TM, Rescorla LA. (2000). Manual for the ASEBA Preschool Forms & Profiles. Burlington, VT: University of Vermont, Research Center for Children, Youth, & Families.

195. Boberg E, Kully D. (1994). Long-term results of an intensive treatment program for adults and adolescents who stutter. J Speech Lang Hear R, 37(5), 1050-9.

196. Newcombe RG. (1998). Two-sided confidence intervals for the single proportion: comparison of seven methods. Stat Med, 17(8), 857-872.

197. Wilson EB. (1927). Probable inference, the law of succession, and statistical inference. J Am Statist Assoc, 22(158), 209-212.

198. Hayes AF, Krippendorf K. (2007). Answering the call for a standard reliability measure for coding data. Commun Methods Meas, 1(1), 77-89.

199. Hayes AF. Spss-macro for calculating Krippendorff's alpha [Internet]. (2013). Available at: http://www.afhayes.com.

200. Holm S. (1979). A simple sequentially rejective multiple test procedure. Scandinavian journal of statistics, 65-70.

201. Perneger TV. (1998). What's wrong with Bonferroni adjustments. BMJ, 316(7139), 1236-1238.

202. Sander EK. (1961). Reliability of the Iowa Speech Disfluency Test. J Speech Hear Disord, (Suppl 7)21-30.

203. Altman DG. (1991). Some common problems in medical research. Practical statistics for medical research, 1396-403.

204. Imel Z, Wampold BE. (2008). The Importance of treatment and the science of common factors in psychotherapy. In: Brown SD, Lent RW, eds. Handbook of counseling Psychology. 4th ed. New York: John Wiley & Sons Inc.

205. Luborsky L, Singer B, Luborsky L. (1975). Comparative studies of psychotherapies. Is it true that "everywon has one and all must have prizes"? Arch Gen Psychiatry, 32(8), 995-1008.

206. Bernstein Ratner N. (2005). Evidence-based practice in stuttering: Some questions to consider. J Fluency Disord, 30(3), 163-188.

207. Hayhow R. (2011). Does it work? Why does it work? Reconciling difficult questions. Int J Lang Commun Disord, 46(2), 155-168.

208. Donaghy M, Harrison E, O'Brian S, Menzies R, Onslow M, Packman A, et al. (2015). An investigation of the role of parental request for self-correction of stuttering in the Lidcombe Program. Int J Speech Lang Pathol, 1-7.

209. Bernstein Ratner N, Guitar B. (2006). Treatment of Very Early Stuttering and Parent-Administered Therapy: the State of the Art. In: Bernstein Ratner N, Tetnowski JA, eds. Current Issues in Stuttering Research and Practice. 1st ed. New York: Psychology Press.

210. Robey RR. (2004). A five-phase model for clinical-outcome research. J Commun Disord, 37(5), 401-411.

211. Conture EG. (1996). Treatment efficacy: stuttering. J Speech Hear Res, 39(5), S18-26.

212. Blumgart E, Tran Y, Craig A. (2010). An investigation into the personal financial costs associated with stuttering. J Fluency Disord, 35(3), 203-215.

213. de Sonneville-Koedoot C, Stolk E, Rietveld T, Franken MC. (2015). Direct versus Indirect Treatment for Preschool Children who Stutter: The RESTART Randomized Trial. PLoS One, 10(7), e0133758.

214. Arnott S, Onslow M, O'Brian S, Packman A, Jones M, Block S. (2014). Group lidcombe program treatment for early stuttering: a randomized controlled trial. J Speech Lang Hear Res, 57(5), 1606-1618.

215. Parkin D, Devlin N. (2006). Is there a case for using visual analogue scale valuations in cost-utility analysis? Health Econ, 15(7), 653-664.

216. College voor zorgverzekeringen (CVZ). (2006). Guidelines for Pharmacoeconomic Research, updated version. Diemen: CVZ.

217. Bex, P. H. H., van der Poll, P. A. M., Bovens, F., Marinkovic, D., & Goo, F. S. (2011). Kostprijsonderzoek logopedie: Onderzoek naar de kosten, inkomsten en tijdsbestedingen van logopediepraktijken. SIRA Consulting.

218. Statistics Netherlands. (2014). Consumentenprijzen; inflatie vanaf 1963. Retrieved from: www.cbs.nl.

R

219. Hakkaart-van Roijen L, Tan S, Bouwmans C. (2010). Handleiding voor kostenonderzoek. Methoden en standaard kostprijzen voor economische evaluaties in de gezondheidszorg. Geactualiseerde versie. College voor Zorgverzekeringen.

220. Ramsey S, Willke R, Briggs A, Brown R, Buxton M, Chawla A, et al. (2005). Good research practices for cost-effectiveness analysis alongside clinical trials: the ISPOR RCT-CEA Task Force report. Value in Health, 8(5), 521.

221. Rubin DB, Schenker N. (1991). Multiple imputation in healthcare databases: An overview and some applications. Stat Med, 10(4), 585-598.

222. Briggs A, Clark T, Wolstenholme J, Clarke P. (2003). Missing.... presumed at random: cost-analysis of incomplete data. Health Econ, 12(5), 377-392.

223. Schafer J. (1997). Analysis of Incomplete Multivariate Data. Number 72 in Monographs on Statistics and Applied Probability. New York: Chapman & Hall.

224. Oostenbrink JB, Al MJ. (2005). The analysis of incomplete cost data due to dropout. Health Econ, 14(8), 763-776.

225. Lewis G, Peake M, Aultman R, Gyldmark M, Morlotti L, Creeden J, et al. (2010). Cost-effectiveness of erlotinib versus docetaxel for second-line treatment of advanced non-small-cell lung cancer in the United Kingdom. J Int Med Res, 38(1), 9-21.

226. de Sonneville-Koedoot C, Stolk EA, Raat H, Bouwmans-Frijters C, Franken MC. (2014). Health-related quality of life of preschool children who stutter. J Fluency Disord, 42, 1-12.

227. Maguire GA, Yeh CY, Ito BS. (2012). Overview of the Diagnosis and Treatment of Stuttering. Journal of Experimental & Clinical Medicine, 4(2), 92-97.

228. Rothwell PM. (2005). External validity of randomised controlled trials: "to whom do the results of this trial apply?". Lancet, 365(9453), 82-93.

229. Green J, Britten N. (1998). Qualitative research and evidence based medicine. BMJ, 316(7139), 1230-1232.

230. Harrison E, Ttofari K, Rousseau I, Andrews C. (2003). Troubleshooting. In: M. Onslow, A. Packman & E. Harrison. The Lidcombe program of early stuttering intervention: A clinician's guide (pp. 91-102). Austin, Texas: Pro-ed.

231. Rousseau I, Packman A, Onslow M, Dredge R, Harrison E. (2002). Australian speech pathologists' use of the Lidcombe Program of early stuttering intervention. Acquiring Knowledge in Speech, Language and Hearing, 4(2), 67-71.

232. O'Brian S, Iverach L, Jones M, Onslow M, Packman A, Menzies R. (2013). Effectiveness of the Lidcombe Program for early stuttering in Australian community clinics. Int J Speech Lang Pathol, 15(6), 593-603.

233. Rogers Everett M. (1995). Diffusion of innovations. New York: Free Press.

234. Kitzinger J. (1994). The methodology of focus groups: the importance of interaction between research participants. Sociol Health Illn, 16(1), 103-121.

235. Kitzinger J. (1995). Qualitative research: introducing focus groups. BMJ, 311(7000), 299.

236. Shek DT, Sun RC, Tang CY. (2009). Focus group evaluation from the perspective of program implementers: findings based on the secondary 2 program. ScientificWorldJournal, 9992-1002.

237. Creswell JW. (2003). Research design. California, CA: Sage Thousand Oaks.

238. Eggers K, De Nil LF, Van den Bergh BR. (2012). The efficiency of attentional networks in children who stutter. J Speech Lang Hear Res, 55(3), 946-959.

239. Jones M, Onslow M, Harrison E, Packman A. (2000). Treating stuttering in young children: predicting treatment time in the Lidcombe Program. J Speech Lang Hear Res, 43(6), 1440-1450.

240. Kingston M, Huber A, Onslow M, Jones M, Packman A. (2003). Predicting treatment time with the Lidcombe Program: replication and meta-analysis. Int J Lang Commun Disord, 38(2), 165-177.

241. Natke U, Sandrieser P, Pietrowsky R, Kalveram KT. (2006). Disfluency data of German preschool children who stutter and comparison children. J Fluency Disord, 31(3), 165-176.

242. Wingate ME. (2001). SLD is not stuttering. J Speech Lang Hear Res, 44(2), 381-383.

243. Cordes AK, Ingham RJ. (1995). Judgments of stuttered and nonstuttered intervals by recognized authorities in stuttering research. J Speech Hear Res, 38(1), 33-41.

244. Howell P. (2013). Screening school-aged children for risk of stuttering. J Fluency Disord, 38(2), 102-123.

245. Gomersall T, Spencer S, Basarir H, Tsuchiya A, Clegg J, Sutton A, et al. (2015). Measuring quality of life in children with speech and language difficulties: a systematic review of existing approaches. Int J Lang Commun Disord, 50(4), 416-435.

246. De Civita M, Regier D, Alamgir AH, Anis AH, Fitzgerald MJ, Marra CA. (2005). Evaluating health-related quality-of-life studies in paediatric populations: some conceptual, methodological and developmental considerations and recent applications. Pharmacoeconomics, 23(7), 659-685.

247. Upton P, Lawford J, Eiser C. (2008). Parent-child agreement across child health-related quality of life instruments: a review of the literature. Qual Life Res, 17(6), 895-913.

248. Griebsch I, Coast J, Brown J. (2005). Quality-adjusted life-years lack quality in pediatric care: a critical review of published cost-utility studies in child health. Pediatrics, 115(5), e600-14.

249. Petrou S. (2003). Methodological issues raised by preference-based approaches to measuring the health status of children. Health Econ, 12(8), 697-702.

250. Dolan P, Kahneman D. (2008). Interpretations of utility and their implications for the valuation of health. The Economic Journal, 118(525), 215-234.

251. Gabel RM, Blood GW, Tellis GM, Althouse MT. (2004). Measuring role entrapment of people who stutter. J Fluency Disord, 29(1), 27-49.

252. Logan KJ, O'Connor EM. (2012). Factors affecting occupational advice for speakers who do and do not stutter. J Fluency Disord, 37(1), 25-41.

253. Dolan P. (2013). Addressing misconceptions in valuing health. Expert Rev Pharmacoecon Outcomes Res, 13(1), 1-3.

254. Stolk EA, Busschbach JJ, Caffa M, Meuleman EJ, Rutten FF. (2000). Cost utility analysis of sildenafil compared with papaverine-phentolamine injections. BMJ, 320(7243), 1165-1168.

255. Stolk EA, Brouwer WB, Busschbach JJ. (2002). Rationalising rationing: economic and other considerations in the debate about funding of Viagra. Health Policy, 59(1), 53-63.

256. Al MJ, Koopmanschap MA, van Enckevort PJ, Geertsma A, van der Bij W, de Boer WJ, et al. (1998). Cost-effectiveness of lung transplantation in The Netherlands: a scenario analysis. Chest, 113(1), 124-130.

257. Hurwitz B. (1999). Legal and political considerations of clinical practice guidelines. BMJ, 318(7184), 661-664.

R