From the Department of Women's and Children's Health
Karolinska Institutet, Stockholm, Sweden

RARE PEDIATRIC MOVEMENT DISORDERS: CLINICAL ASPECTS OF GENOTYPE, PHENOTYPE, AND ASSESSMENT

Annika Danielsson

Karolinska Institutet

Stockholm 2022

Published by Karolinska Institutet.
Printed by Universitetsservice US-AB, 2022

ISBN 978-91-8016-694-2

Rare Pediatric Movement Disorders: Clinical Aspects of Genotype, Phenotype, and Assessment

THESIS FOR DOCTORAL DEGREE (Ph.D.)

By

Annika Danielsson

Public defense in Skandiasalen, 1st floor, Building Q1, Karolinska University Hospital, Solna, Friday, November 18, 2022, at 9:00 am

Principal Supervisor:
Associate Professor Kristina Tedroff
Karolinska Institutet
Department of Women's and Children's Health
Division of Neuropediatrics

Co-supervisors:
Associate Professor Lena Krumlinde-Sundholm
Karolinska Institutet
Department of Women's and Children's Health
Division of Neuropediatrics

Professor Ann Nordgren
Karolinska Institutet
Department of Molecular Medicine and Surgery
Division of Rare Diseases

Opponent:
Professor Jonathan W. Mink
University of Rochester
Departments of Neurology, Neuroscience and Pediatrics

Examination Board:
Associate Professor Kate Himmelmann
University of Gothenburg
Institute of Clinical Sciences
Department of Pediatrics

Associate Professor Gustav Jarl
Örebro University
Faculty of Medicine and Health
School of Health Sciences

Associate Professor Rickard L Sjöberg
Umeå University
Department of Clinical Sciences
Division of Neurosciences

This thesis is dedicated to all participating children and young adults, as well as their families, for their contributions to this work and for letting others benefit from these findings in the future.

POPULAR SCIENCE SUMMARY OF THE THESIS

This thesis explores clinical aspects of rare pediatric movement disorders and is focused on genetic causes, clinical pictures (genotype, phenotype), and assessments. Movement disorders are a group of disorders with impaired ability to initiate and control movements. Children with movement disorders often suffer from hyperkinetic movements, with unwanted and extra movements, that disturb the performance of voluntary movements and the maintenance of a stable posture. Dystonia and choreoathetosis are two of those hyperkinetic movements. Pediatric movement disorders with dystonia and choreoathetosis are rare, and our knowledge regarding prognosis and treatment options is limited.

During the past decades, however, our understanding of these rare pediatric movement disorders has increased. Progress in the genetic field has had great impact on children receiving a causative diagnosis, and consequently, our knowledge regarding genotype–phenotype correlation in different conditions is increasing. Consensus has been reached regarding definitions and classification of hyperkinetic movements in childhood. Assessment scales have been developed to evaluate the presence of movement disorders in children before and after interventions. Deep brain stimulation (DBS), a neurosurgical treatment, has turned out to be an effective treatment option for some, but not all, individuals with dystonia.

The overall aim of this thesis was to enable and facilitate assessment of dystonia and choreoathetosis in children and young adults, and to elucidate some clinical aspects in two rare pediatric-onset dystonia disorders.

The Dyskinesia Impairment Scale (DIS) is an assessment scale designed for evaluation of dystonia and choreoathetosis in children and young adults with dyskinetic cerebral palsy. In Study III, we examine psychometric properties of the DIS in a new population; children and young adults with inherited dystonia and choreoathetosis. We have been able to show that the DIS is reliable and valid in these individuals as well. Altogether, these results indicate that the DIS might be a valuable tool to evaluate children and young adults with inherited dystonia either longitudinally, or before and after interventions.

The DIS is a comprehensive assessment scale, and it takes a long time to administer and to score. In Study IV, we develop a shortened and more user-friendly version of the DIS, the DIS-II. Scale development included an online advisory expert meeting, iterative discussions within the research team, and a Rasch measurement model analysis. The Rasch analysis is a method that is widely being used to create new test instruments or to evaluate existing scales. The DIS-II consists of 60% of the items in the original DIS, with a reduced number of scoring steps for each item. The Rasch analysis demonstrated good to excellent validity and reliability, implying that DIS-II can be used in the clinic as well as in research when assessing dystonia and choreoathetosis.

Benign paroxysmal torticollis of infancy (BPTI) is a rare pediatric movement disorder where affected children suffer from recurrent episodes of dystonia in the neck region. BPTI is self-limiting and has traditionally been referred to as a benign condition. However, genetic research has linked BPTI to less benign disorders and delayed motor development has been described in BPTI. Study I, a long-term follow-up study, indicates that despite this, BPTI does not lead to negative lasting consequences. At the last follow-up, with a mean age of 14 years, all children

attended regular school, and no one was described as having motor problems. Five out of 11 had developed migraine and/or episodic syndromes that may be associated with migraine, but no one had been diagnosed with a more severe disorder. We found no known mutations in the genes we explored.

DYT-THAP1 dystonia is a rare pediatric-onset movement disorder, where affected individuals suffer from a severe, life-long motor disability. Commonly, the dystonia is first noticeable in the head and neck region, before generalizing to the rest of the body. Speech and swallowing are often impaired. Study II could confirm that DBS is an effective treatment option for these individuals. Fourteen individuals were followed for a median time of nearly five years, and at the last follow-up, median dystonia reduction was about 60% when assessed with the Burke–Fahn–Marsden Movement Scale (BFM-M). However, improvement was less obvious regarding speech and swallowing.

ABSTRACT

Pediatric movement disorders are defined by an impaired ability to initiate and control movements. Dystonia, one of these disorders, may cause a life-long severe motor disability. Pharmacological treatment is often insufficient, but deep brain stimulation (DBS) has turned out to be effective for some dystonia subtypes. Reliable and valid assessment tools are needed to evaluate new interventions for dystonia. The Dyskinesia Impairment Scale (DIS) is an assessment scale for dystonia and choreoathetosis, designed for children and youth with dyskinetic cerebral palsy. However, the DIS is comprehensive, and its administration and scoring times are long, and the scale is often disregarded in a clinical setting.

The overall aim of this thesis was to enable and facilitate assessment of dystonia and choreoathetosis in children and young adults, and to elucidate some clinical aspects in two rare pediatric-onset dystonia disorders.

Study I: This prospective long-term case series indicates that benign paroxysmal torticollis of infancy does not lead to neurological sequelae. At the last follow-up, with a mean age of 14 years, all 11 children attended regular school, and no one reported motor problems. Five had developed migraine and/or episodic syndromes that may be associated with migraine. We found no known mutations in candidate genes. Nevertheless, in the case of a severe phenotype, specifically if there are paroxysmal disorders in the family history, it may be important to follow the children longitudinally and genetic testing can be considered.

Study II: This retrospective multi-center case series confirms that pallidal DBS is an effective treatment option for individuals with DYT-THAP1 dystonia. Fourteen individuals were followed for a median time of nearly five years, and at the last follow-up, median dystonia reduction was about 60% when assessed with the Burke–Fahn–Marsden Movement Scale (BFM-M). However, improvement was less obvious in the orolaryngeal and craniocervical regions; anatomical areas which are often affected by dystonia in these individuals.

Study III: This cross-sectional two-center study was an instrument evaluation of the DIS for a new population; children and young adults with inherited or idiopathic dystonia. The DIS and the dystonia subscale were shown to have good-to-excellent inter-rater and test–retest reliability, whereas the choreoathetosis subscale had moderate inter-rater and excellent test–retest reliability. Concurrent validity for the dystonia subscale, when compared with the gold standard BFM-M, was good. Altogether, these results indicate that the DIS might be a valuable tool to evaluate dystonia and choreoathetosis in inherited and idiopathic dystonia.

Study IV: This cross-sectional two-center study aimed to design a more clinically useful version of the DIS, the DIS-II. Scale development included an online advisory expert meeting, iterative discussions within the research team, and a Rasch measurement model analysis. With a reduced number of scoring steps, the DIS-II evaluates 60% of the items in the original DIS. The result from the Rasch analysis demonstrates evidence of construct validity of the scale. Furthermore, a high person reliability indicates that the DIS-II may separate assessed individuals into eight distinct ability levels. Altogether, this implies that the DIS-II provides valid and reliable measures for dystonia and choreoathetosis and reduces administration and scoring times in comparison with the DIS.

LIST OF SCIENTIFIC PAPERS

I. Annika Danielsson, Britt-Marie Anderlid, Tommy Stödberg, Kristina Lagerstedt-Robinson, Eva Klackenberg Arrhenius, Kristina Tedroff
Benign paroxysmal torticollis of infancy does not lead to neurological sequele
Dev Med Child Neurol. 2018 Dec;60(12):1251-1255. doi: 10.1111/dmcn.13939. Epub 2018 Jun 28.

II. Annika Danielsson, Miryam Carecchio, Laura Cif, Anne Koy, Jean-Pierre Lin, Göran Solders, Luigi Romito, Katja Lohmann, Barbara Garavaglia, Chiara Reale, Giovanna Zorzi, Nardo Nardocci, Philippe Coubes, Victoria Gonzalez, Agathe Roubertie, Gwenaelle Collod-Beroud, Göran Lind, Kristina Tedroff
Pallidal Deep Brain Stimulation in DYT6 Dystonia: Clinical Outcome and Predictive Factors for Motor Improvement
J Clin Med. 2019 Dec 6;8(12):2163. doi: 10.3390/jcm8122163

III. Annika Danielsson*, Inti Vanmechelen*, Cecilia Lidbeck, Lena Krumlinde-Sundholm, Els Ortibus, Elegast Monbaliu, Kristina Tedroff
*these authors contributed equally to the study
Reliability and Validity of the Dyskinesia Impairment Scale in Children and Young Adults with Inherited or Idiopathic Dystonia
J Clin Med. 2020 Aug 11;9(8):2597. doi: 10.3390/jcm9082597.

IV. Inti Vanmechelen*, Annika Danielsson*, Cecilia Lidbeck, Kristina Tedroff, Elegast Monbaliu, Lena Krumlinde-Sundholm
*these authors contributed equally to the study
The Dyskinesia Impairment Scale-II: Development, Construct Validity and Reliability *Manuscript, accepted for publication in Developmental Medicine & Child Neurology*

CONTENTS

1 INTRODUCTION....5
1.1 Rare disorders....5
1.1.1 Definition....5
1.1.2 Rationale for studying rare disorders....5
1.1.3 Special concerns when studying rare disorders....6
1.2 Movement disorders....6
1.2.1 Pediatric movement disorders....7
1.2.2 Rationale for studying pediatric movement disorders....7
1.2.3 Benign paroxysmal torticollis of infancy (BPTI)....8
1.2.4 DYT-THAP1 dystonia....8
1.3 Dystonia....8
1.3.1 Clinical features and classification of dystonia....9
1.3.2 Treatment of pediatric dystonia....10
1.3.3 Assessment of dystonia....11
1.4 Measurement properties....13
1.4.1 Reliability....13
1.4.2 Validity....14
1.4.3 Theories of measurement....15
2 AIMS OF THE THESIS....17
3 METHODS....19
3.1 Study designs....19
3.2 Participants....19
3.2.1 Recruitment....20
3.2.2 Sample size....20
3.3 Procedure and data collection....21
3.3.1 Study I....21
3.3.2 Study II....21
3.3.3 Study III....22
3.3.4 Study IV....22
3.4 Genetic analysis....23
3.4.1 Sanger sequencing....23
3.4.2 Whole-exome sequencing....24
3.5 Assessment scales....24
3.5.1 Peabody Developmental Motor scale-2....24
3.5.2 Movement Assessment Battery for Children....24
3.5.3 The Dyskinesia Impairment Scale (DIS)....24
3.5.4 The Burke–Fahn–Marsden Dystonia Rating Scale (BFM)....26
3.6 Statistical methods....26
3.6.1 Descriptive statistics....27
3.6.2 Comparing groups and examining correlations....27
3.6.3 Measures of reliability....28

3.6.4 Measures of validity 29
3.6.5 Rasch measurement model analysis 29
3.7 Ethical considerations 33
4 RESULTS AND DISCUSSION 35
4.1 Study I: Genotype, phenotype, and long-term development in BPTI 35
4.1.1 Nature of episodes 36
4.1.2 Family history 36
4.1.3 Development of paroxysmal disorders 37
4.1.4 Genetic findings 37
4.1.5 Long-term development: motor function and cognitive abilities 38
4.1.6 Strengths and limitations 39
4.2 Study II: Response to Deep Brain Stimulation in DYT-THAP1dystonia 39
4.2.1 Aspects of genotype 41
4.2.2 Aspects of phenotype 41
4.2.3 Aspects of location of leads and stimulation parameters 42
4.2.4 Aspects of timing of surgery 43
4.2.5 Strengths and limitations 43
4.3 Study III: Reliability and validity of the DIS in inherited dystonia 44
4.3.1 Inter-rater reliability 45
4.3.2 Test-retest reliability 47
4.3.3 Concurrent validity 48
4.3.4 Advantages and disadvantages of the DIS 48
4.3.5 Strengths and limitations 49
4.4 Study IV: The DIS-II: Development, construct validity, and reliability 50
4.4.1 Development of the DIS-II 50
4.4.2 Rasch analysis of the DIS-II 52
4.4.3 Advantages and disadvantages of the DIS-II 54
4.4.4 Strengths and limitations 55
5 CONCLUSIONS AND CLINICAL IMPLICATIONS 57
6 FUTURE PERSPECTIVES 59
7 ACKNOWLEDGEMENTS 61
8 REFERENCES 63

LIST OF ABBREVIATIONS

BADS	Barry–Albright Dystonia Scale
BFM	Burke–Fahn–Marsden Dystonia Rating Scale
BFM-D	BFM Disability scale
BFM-M	BFM Movement scale
BPTI	Benign Paroxysmal Torticollis of Infancy
BPV	Benign Paroxysmal Vertigo
CP	Cerebral Palsy
CTT	Classical Test Theory
DBS	Deep Brain Stimulation
DIS	Dyskinesia Impairment Scale
DIS-CA	Choreoathetosis Subscale of the DIS
DIS-D	Dystonia Subscale of the DIS
DIS-II	Dyskinesia Impairment Scale-II
DIS-II-CA	Choreoathetosis Subscale of the DIS-II
DIS-II-D	Dystonia Subscale of the DIS-II
DNA	Deoxyribonucleic Acid
dNTP/ddNTP	Deoxynucleotide/dideoxynucleotide triphosphates
DYT	Dystonia nomenclature for monogenic dystonia disorders
GDS	Global Dystonia Rating Scale
GPi	Globus Pallidus interna
ICC	Intraclass Correlation Coefficient
ICF	International Classification of Functioning, Disability and Health
ICHD-3	International Classification of Headache Disorders 3rd edition
IQR	Interquartile range
IRT	Item Response Theory
ITB	Intrathecal Baclofen
MDD	Minimal Detectable Difference
MnSq	Mean Square
PCA	Principal Component Analysis

SD	Standard Deviation
SEM	Standard Error of Measurement
STN	Subthalamic nucleus
UDRS	Unified Dystonia Rating Scale
VIM	Ventralis Intermedius nucleus of the Thalamus
WES	Whole Exome Sequencing
WGS	Whole Genome Sequencing

1 INTRODUCTION

Pediatric movement disorders are a heterogeneous group of disorders defined by an impaired ability to initiate and control movements. In dystonia, one of these conditions, affected individuals often suffer from a life-long severe motor disability. The etiology and clinical picture of dystonia is highly variable, and many of these disorders are rare. Diagnostic work-up is often challenging, and it is very important for a child to receive an early diagnosis, since some underlying disorders are treatable. For others, a correct etiologic diagnosis may help families manage expectations regarding prognosis or treatment options.

Pharmacological treatment of dystonia is often disappointing, but in the last decades, deep brain stimulation (DBS) has turned out to be an effective treatment option for some, but not all, affected individuals. The variable response to DBS is not fully understood, but etiology of the dystonia explains some of the variation. To select patients for surgery, more knowledge about DBS outcome in different dystonia subtypes is needed.

Reliable and valid assessment tools are needed to evaluate new interventions for dystonia. The Dyskinesia Impairment Scale (DIS) is an assessment scale for dystonia and choreoathetosis, designed for children and youth with dyskinetic cerebral palsy (CP). This scale has shown good to excellent psychometric properties, indicating that it is a valuable tool in intervention studies. However, the DIS takes a long time to administer and score and is therefore not often used in a clinical setting.

1.1 RARE DISORDERS

1.1.1 Definition

There is no worldwide accepted definition for a rare disease or disorder. In Sweden, as well as in the rest of Europe, a rare disease is defined as affecting less than 5 in 10,000 individuals(1, 2). In the United States though, with a population of about 335 million people, a disease is defined as ‘rare’ when it affects fewer than 200,000 individuals living in the States at any given time(3, 4). Most rare diseases are inherited, although some are caused by cancers, infections, or rare autoimmune disorders(5).

1.1.2 Rationale for studying rare disorders

The number of known rare diseases is increasing, and today more than 6,000 different forms of rare disorders have been described(5, 6). Although rare, since there are so many of these disorders, they are actually affecting a great number of individuals. As much as about 3%–6% of the worldwide population is currently suffering from a rare disorder(7). Individuals suffering from rare disorders often share some common features. First, there is often a delay in the diagnosis. Second, knowledge of the symptoms and the prognosis of the disease is often lacking. Furthermore, many of the conditions are severe, and may lead to a shortened life span. Additionally, since a majority of the diseases are genetic, many families are interested in

genetic counseling. Lastly, and perhaps most importantly, pharmacological treatment options are few, and often used off-label. Hence, the rationale for studying rare disorders is obvious; as a group the disorders are common, and there are many important unanswered questions to address.

1.1.3 Special concerns when studying rare disorders

When studying rare disorders, there are some common challenges that one encounters as a researcher. Small sample sizes are inevitable, since recruitment of study participants is not straightforward. A small sample size has a higher risk of not being representative of the population of interest. Heterogeneity, i.e. diversity in the characteristic you are studying, is often large in small samples, and heterogeneity increases the risk of finding a false statistically significant difference. Collaboration with other research centers, and using networks for rare disorders and patient advocacy groups, can be helpful in the recruitment process. Today the internet may also be useful, where you as a researcher may advertise on Twitter or Facebook pages to reach out to support groups of rare disorders.

Another challenge is to protect the privacy of the study subjects, as the publication of family trees might lead to identification of a specific individual. Furthermore, research funding is more difficult to obtain compared to the study of common disorders, as common disorders have greater economic impact.

From the perspective of study design, different options are possible. When it comes to observational studies, case reports or small case series are sometimes the only option. Case-control studies have a higher level of evidence compared to case reports or small case series, and are often a good choice when studying rare disorders. Persons with a rare disorder are compared with suitable persons without the disorder (controls), and the relationship between the odds of having a specific risk factor/exposure in the two groups is examined (odds ratio). When it comes to experimental studies, a cross-over design might be an option, where the person serves as their own control.

1.2 MOVEMENT DISORDERS

Movement disorders are defined by the presence of involuntary movements, or by impaired ability to perform voluntary movements, or to maintain a stable posture(8-10). These abnormal movements are not caused by weakness, and muscle tone may be increased or decreased. The pathophysiology of movement disorders is not fully understood, but damages or disorders affecting neural circuits in subcortical or cerebellar structures are known to cause movement disorders(11). These neural circuits are involved in learning, planning, and inhibition of movements, and precede the final motor pathway from the primary motor cortex. Movement disorders can broadly be classified as either hyperkinetic, with unwanted excess movements, or as hypokinetic, with too little movement. In some cases, the abnormal movement is the only symptom, but often it is combined with other neurological problems.

1.2.1 Pediatric movement disorders

Movement disorders affecting children differ from those seen in adulthood(12, 13). Normal motor patterns, as well as dysfunctional motor symptoms, are changing during the period when the brain and nervous system develops and matures, and specific movement disorders tend to develop at certain ages. Pediatric movement disorders can be transient, but this very rarely occurs in adulthood(14). Furthermore, children often suffer from hyperkinetic movements, such as tics, tremor, dystonia, and choreoathetosis, whereas adults more often have hypokinetic movement disorders, such as Parkinson's disease(15).

In this project, the focus is on pediatric hyperkinetic movement disorders, in particular dystonia, and to a lesser extent, chorea and athetosis. Dystonia, chorea, and athetosis often co-occur, as in dyskinetic CP(16, 17). The definitions of these movement disorders, according to the Taskforce on Childhood Movement disorders are as follows: "*Dystonia* is a movement disorder in which involuntary sustained or intermittent muscle contractions cause twisting and repetitive movements, abnormal postures, or both"(15). "*Chorea* is an ongoing, random-appearing sequence of one or more discrete involuntary movement or movement fragments"(15). "*Athetosis* is a slow, continuous, involuntary writhing movement that prevents maintenance of a stable posture"(15).

1.2.2 Rationale for studying pediatric movement disorders

Pediatricians often meet parents who are worried about their children having unusual movement patterns. In some of these cases, the disorders are benign and transient, and families just need information about the clinical course and prognosis(14). In other cases, the movement disorder severely affects the children, causing activity limitation and impairing their quality of life(18, 19). Some pediatric movement disorders are common and well known, for example cerebral palsy (CP) and tics, but many disorders are rare, each affecting only a few individuals. Regarding rare pediatric movement disorders, it may initially be difficult to distinguish between benign and transient variants, and those that are severe and persistent. Diagnostic delay is common, and many children will never receive an exact diagnose(20). It is not uncommon for a rare genetic or inflammatory movement disorder—sometimes one with severe motor symptoms—to be mistaken for a psychiatric or functional diagnosis.

Today, the chance of a child receiving an accurate diagnosis and disease-specific or symptom-based treatment, is probably higher than just a couple of years ago. During the past two decades, pediatric movement disorders, as a field, has been rapidly expanding(21, 22). Advances in neurogenetics, with the use of next-generation sequencing, has facilitated the identification of causative genes, and this has led to a better understanding of the phenotypic spectrum of many conditions with movement disorders(23-25). The role of neuroimmunology in some movement disorders has been highlighted, which has had crucial implications regarding treatment(26, 27). Deep brain stimulation (DBS), an invasive neurosurgical intervention, has been shown to be very efficient in treating different kinds of movement disorders, including dystonia(28, 29). New consensus definitions for childhood movement disorders have been accepted, which has had a positive impact on research(15, 30, 31). But despite these advances, there is still much more to explore.

1.2.3 Benign paroxysmal torticollis of infancy (BPTI)

Benign paroxysmal torticollis of infancy (BPTI) is a rare transient pediatric movement disorder, with recurrent, stereotypic episodes of cervical dystonia with disease onset in early infancy(32-34). The episodes occur regularly, typically once a month, and last a couple of days. The torticollis is often combined with irritability, pallor, vomiting, and ataxia. The children start to improve at the age of two to three years, and by the age of five, the torticollis episodes have normally disappeared. The quality of life in infants suffering from BPTI has been reported to be significantly affected(35). Motor delay has been reported in some children(34, 36). Neuroradiology, EEG and laboratory tests are usually normal(34). Only case reports or small case series have been published, and the etiology and long-term development of BPTI is not fully known.

There is an association between BPTI and migraine. A family history of migraine is common(33, 34, 37). Migraine or other syndromes associated with migraine, for example benign paroxysmal vertigo, cyclic vomiting, or abdominal migraine, are overrepresented in individuals who have had BPTI(37-39). Mutations in *CACNA1A* and *PRRT2* have been described in BPTI(38, 40-43). Such mutations, as well as mutations in *ATP1A2* and *SCN1A*, are known to cause familial hemiplegic migraine. Mutations in these four genes are also described in epilepsy, episodic ataxia, paroxysmal tonic upgaze, alternating hemiplegia of childhood, and paroxysmal dyskinesia(23, 44-47). BPTI has been regarded as a benign condition, but considering reports regarding motor delay and new knowledge about genetic connections to non-benign conditions, this needs to be confirmed. Before the publication of Study I in this thesis, long-term follow-up studies on BPTI were lacking.

1.2.4 DYT-THAP1 dystonia

DYT-THAP1 dystonia is a rare pediatric-onset persistent dystonia disorder caused by mutations in the *THAP1* gene(48, 49). The disorder is characterized by isolated dystonia, where cranio-cervical, laryngeal and oromandibular dystonia are frequently observed at disease onset, and generalization is common. Pharmacological efforts to alleviate symptoms have so far been disappointing(50). Deep Brain Stimulation (DBS) is known to reduce the amount of dystonia very effectively in some individuals with isolated dystonia(51-53). Limited and inconsistent data exist about the outcome after DBS in DYT-THAP1 dystonia(49, 54-57). Therefore, there is a need for studies exploring outcomes after DBS in DYT-THAP1 dystonia.

1.3 DYSTONIA

The word “dystonia” was first used by Oppenheim in 1911, when he reported on a disease with the proposed name of “dystonia musculorum deformans”(58). He postulated that the disorder had an organic etiology, but for a long time the disorder was believed to be psychogenic. In 1997, the DYT1-gene (TOR1A) was described as the cause of Oppenheim’s primary torsion dystonia(59). Today we know that some conditions with dystonia are inherited, whereas others are acquired; for example due to perinatal asphyxia, infections, neoplastic diseases, or traumatic brain injuries(15, 30, 60). Dystonia is the second-most common pediatric movement

disorder (after tic disorder), and dyskinetic CP is the most common disorder with pediatric-onset dystonia(19, 20). Since pediatric-onset dystonia is unusual and heterogenous, there are only few reports of prevalence, natural history, and quality of life of children with dystonia(19, 61-63).

1.3.1 Clinical features and classification of dystonia

Dystonia is characterized by involuntary sustained or intermittent muscle contractions, causing twisting and repetitive movements, abnormal postures, or both(15). Dystonic muscle contractions are typically triggered by voluntary movement, and during sleep they are often absent. Dystonia is predictable, since each individual has a characteristic pattern of repeated dystonic movements and postures. Some specific postures are shared by many, such as blepharospasm or foot inversion. Task specificity is a special feature of dystonia. Dystonic movements are elicited by particular tasks, such as writing or playing an instrument, whereas other movements, involving the same muscles, do not. Sensory trick is another interesting feature, where the dystonia may (at least shortly) be reduced by a gentle touch on the skin near the site of dystonia. A slight touch on the chin may cause partial relief of cervical dystonia.

Dystonia may be focal, segmental or generalized(15, 30, 60). The onset of dystonia can be in childhood or in adulthood and it might be insidious or rapid. Dystonia may be isolated or combined with other movement disorders, such as spasticity, rigidity, choreoathetosis, myoclonus, and tremor(15, 30, 60). Other neurological or systemic disorders may also co-occur with dystonia. Thus, clinical features range from isolated adult-onset focal dystonia to severe childhood-onset generalized dystonia combined with other neurological symptoms.

Dystonia has traditionally been classified as ‘primary’ or ‘secondary’. ‘Primary’, in this setting, means that the dystonia is isolated and is not combined with other movement disorders or neurological symptoms. Primary dystonia also implicates that they are inherited and that there are no signs of structural brain damage. The term secondary dystonia is less clear, sometimes indicating non-primary dystonia, sometimes indicating the dystonia has a defined pathology, is acquired, or has a known cause. ‘Heredodegenerative disorders’ is a terminology used to address secondary dystonias caused by metabolic disorders. However, as our understanding of dystonia disorders and their genetic causes has increased, this classification is inconsistent.

In 2013, a consensus group developed a new classification system for dystonia(60). This classification system consists of two distinct axes: clinical characteristics (age at onset, body distribution, temporal pattern, associated features) and etiology (anatomical changes, inheritance pattern)(60). The first axis, clinical charateristics, helps the clinician to properly describe the individual phenotype, which facilitates diagnostic work-up and treatment. The second axis, etiology, is an area where knowledge is expanding , and information here need to be updated on a regular basis. The two characteristics of the second axis, anatomical changes and inheritance, should be characterized separately, since anatomical abnormalities might exist in inherited cases. If the dystonia is inherited, you should define the inheritance pattern; autosomal dominant, autosomal recessive, X-linked, or mitochondrial. Although this dystonia classification was designed for an adult population, it can be used for children as well(12).

1.3.2 Treatment of pediatric dystonia

Treatment of children with dystonia differs from that of adults for several reasons. One reason is that pediatric dystonia is often combined with other movement disorders. Other reasons are that the dystonia is influenced by the normal maturation of the brain and that the dystonia is interfering with the development of motor skills. Furthermore, side effects differ between children and adults. There are several reports and reviews published regarding treatment of dystonia in adulthood, but only some for pediatric-onset dystonia(64-68). Pharmacological treatment includes oral medication and botulinum toxin injections(67). For patients where pharmacological therapy is insufficient, there are several surgical procedures, among which intrathecal baclofen and deep brain stimulation are the most important(67). In most patients, the extent of dystonia can be reduced if these methods are combined(66). Despite lacking evidence as to effectiveness, physiotherapy and occupational therapy play an important role in the management of the patients; including, for example, muscle relaxation methods and strengthening of antagonist muscles to balance abnormal postures(66, 69, 70).

When choosing therapy, healthcare providers need to be systematic. First, it is of great importance to conduct a detailed evaluation to rule out disorders that can be treated specifically; for example dopa-responsive dystonia, glucose transport disorders, Wilson's disease or immune-mediated diseases such as NMDA-receptor encephalitis or autoimmune basal ganglia encephalitis(50, 67, 71). Second, other co-occurring motor disorders, such as spasticity, can influence treatment choice. Third, depending on whether the dystonia is focal, segmental, or generalized, different treatment options are possible. Last, comorbidity—for example pain and anxiety—that affects the degree of dystonia, must be considered, and maybe treated.

1.3.2.1 Pharmacological treatment of pediatric dystonia

Generally, there is limited evidence supporting pharmacological treatment of pediatric dystonia, and most patients are treated based on clinical expert opinions(67, 72). *L-dopa* is the first line drug for treatment of pediatric-onset dystonia without a known cause(67, 68). When effective, this can help in identifying individuals with dopa-responsive dystonia, but L-dopa may also be a therapeutic option in other forms of dystonia, for example juvenile Parkinson`s disease or neurodegenerative, metabolic disorders(50, 73). *Trihexyphenidyl* is an anticholinergic drug often used to treat children with generalized dystonia(68, 72). This is the only drug where a randomized trial has documented a positive effect in pediatric dystonia that does not respond to L-dopa treatment(74). In this double-blind randomized placebo-controlled study, 31 children and young adults with segmental and generalized dystonia were treated with high doses of trihexyphenidyl (about 30mg daily), and 71% of them had a clinically significant improvement. *Oral baclofen or benzodiazepines* are options in generalized dystonia combined with spasticity; however, side effects are common(67). There are several other pharmacological options, however evidence is lacking for effect as well as for dosing(61, 66). In focal dystonia, the first step after L-dopa would be intramuscular injections of *botulinum toxin* that effectively and reversibly reduce muscle tone(67).

1.3.2.2 Surgical treatment of pediatric dystonia

Intrathecal baclofen pump (ITB): a pump implanted under the skin delivers baclofen via a subcutaneous catheter into the spinal fluid. When delivered directly to the spinal fluid, the effect of baclofen is improved. The intrathecal administration of baclofen allows reduced doses, and systemic adverse effects can be diminished. In Sweden, ITB is mainly used in non-ambulatory children with dyskinetic cerebral palsy, where dystonia and spasticity are co-occurring.

Deep brain stimulation (DBS): a neurostimulator implanted under the skin sends electrical impulses through subcutaneous electrodes to central brain structures, for example the thalamus or globus pallidus. DBS is a treatment option for dystonia as well as other movement disorders such as essential tremor and Parkinson´s disease(75-78). The mechanism of action is unclear, but it is likely dependent on which target is stimulated and which condition is being treated(79). DBS has been evaluated for a range of different dystonia conditions, and some patients with dystonia benefit more from the procedure than others. The variable response to DBS is not fully understood. In isolated generalized dystonia, DBS produces good or sometimes even excellent results, with motor improvement ranging from 50–75%(51-53). On the other hand, in dystonia combined with other neurological symptoms, or dystonia with pathological findings on MRI, the outcome of DBS is poorer and less predictable(80, 81).

1.3.3 Assessment of dystonia

Evaluating dystonia is difficult for many reasons. First, dystonia is a dynamic condition that changes depending on the posture and on the activity of the involved body area. Non-motor stimuli, such as stress, pain, or sudden noise, may also affect the presence and severity of dystonia. Second, motor symptoms like spasticity, bradykinesia, choreoathetosis, and ataxia may coexist with the dystonia. Additionally, some individuals with dystonia have cognitive impairment. Few assessment tools or scales exist to assess dystonia and most of the existing scales are designed for adults. Application of adult dystonia scales to children is complicated, since typical motor development has to be accounted for. Children are supposed to develop new motor skills over time, thus if the child develops these skills slower than expected or remains at a stable stage in gross motor function, this might indicate a progress of the underlying disorder.

Only a few review articles or original articles evaluate assessment tools for generalized dystonia in adults or children(82-85). For adults, the only recommended scale is the Burke–Fahn–Marsden Dystonia Rating Scale (BFM)(82, 86). For the pediatric dyskinetic cerebral palsy population, the Barry–Albright Dystonia Scale (BADS) is considered most useful for clinical use, and the Dyskinesia Impairment Scale (DIS) for research(83, 87, 88). Today, there is no rating scale that has been designed for children with generalized inherited dystonia. In the absence of recommended scales for these patients, the use of the BFM is widely accepted and used; for example when evaluating children having received DBS.

The BFM, the BADS and the DIS, together with the Unified Dystonia Rating Scale, the Global Dystonia Rating Scale, the Movement Disorder-Childhood Rating Scale, and the Movement Disorder-Childhood Rating Scale 0–3, are currently being used for assessment of children with generalized dystonia.

The Burke–Fahn–Marsden Dystonia Rating Scale (BFM) was developed in 1985 to assess isolated generalized dystonia in adults(86). The BFM consists of a movement subscale (BFM-M), based on examination of the patient, and a disability subscale (BFM-D), based on the patients' or caregivers' report of disability in activities of daily living. BFM-M rates dystonia severity and provoking factors in 9 body areas, including eyes, mouth, speech and swallowing, neck, trunk, and both arms and legs. Additionally, for the mouth, eyes, and neck, there is a weighing factor that halves the input from these body regions. The BFM-M is mainly based on function, and for individuals with cognitive impairment or problems with activities of daily living, the actions included in the assessment may be impossible to perform and instructions may be hard to understand. A video protocol was developed for the BFM-M.

The Unified Dystonia Rating Scale (UDRS) was developed in 1997 for assessment of isolated generalized dystonia in adults(85). This scale was designed to address some perceived short-comings of the BFM-M, including variable definitions of body regions, and the use of weighing factors. Another aim was to create a comprehensive assessment scale. The UDRS evaluates the severity and duration of dystonia in 14 body regions, including eyes and upper face, lower face, jaw and tongue, larynx, neck, trunk, shoulder/proximal arm, and distal leg/foot. The severity rating (0–4) is specific for each body region, whereas the duration factor (0–4) evaluates whether the dystonia occurs in rest, or in action, and whether it is predominantly of a maximal or submaximal intensity. A video protocol was created, with all actions and rest postures of the assessment included.

The Global Dystonia Rating Scale (GDS) was developed together with the UDRS and assesses dystonia in the same 14 body regions as the UDRS(85). Each body region is assessed on a rating scale; 0–10, where 0 indicates no dystonia and 10 the most severe dystonia. Here, no provoking factors or weighing factors are included in the assessment. This scale has been shown to be easy to use, and is therefore recommended for clinical purposes, whereas the UDRS probably is more useful in research(85).

The Barry–Albright Dystonia Scale (BADS) was based on the BFM-M, and was developed in 1999 to assess children and adults before and after receiving intrathecal baclofen, most of them having dyskinetic cerebral palsy(87). The BADS rates 8 body regions including the eyes, mouth, neck, trunk, and four limbs from 0 (no dystonia) to 4 (dystonia present >50% of the time and/or interfering with specified actions). Compared to the BFM-M, the BADS does not assess dystonia during speech and swallowing, and provoking factors are not included in the assessment. The BADS is possible to use in children and in individuals with cognitive impairment who cannot follow or understand instructions.

The Movement Disorder-Childhood Rating Scale and the *Movement Disorder-Childhood Rating Scale 0–3* were developed in 2008 and 2009, respectively, to evaluate different pediatric movement disorders, such as hypokinetic-rigid, chorea/ballismus, dystonia/athetosis, myoclonus, tic, and tremor(89, 90). The scales are divided into two parts; a first part evaluating the degree of disability in different domains (motor function, oral/verbal function, daily living

activities), and a second part with focus on the severity of the most predominant movement disorder for each child. The severity of the movement disorder is evaluated in 7 body regions (eye and periorbital region, face, tongue and perioral region, neck, trunk, upper limbs, and lower limbs). Each body region is scored from no sign of movement disorder (score 0) to a movement disorder that is present in all tasks/prevents the child from performing the tasks (score 4). No provoking or weighing factors are used. Film protocols with actions and rest postures were developed for both scales.

The Dyskinesia Impairment Scale (DIS) was developed 2012 to assess dystonia and choreoathetosis in children and young adults with dyskinetic cerebral palsy (DCP)(88). The DIS evaluates 12 body regions (eyes, mouth, neck, trunk, and proximal/distal limbs) in rest and in action separately. The amount of time (duration) and the range of motion (amplitude) of the dystonia and/or choreoathetosis are assessed on a 5-point rating scale. Specific training in discriminating dystonia and choreoathetosis is probably important before using the DIS. A video protocol with all actions/rest postures exists. The DIS was designed to overcome some limitations perceived with the BADS. First, the DIS also assesses choreoathetosis. Second, the items in the BADS are combinations of several different dystonia characteristics. Third, there is no differentiation between rest and action. Furthermore, the BADS was earlier shown to have a high standard error of measurement(91). Another advantage of the DIS was that it was based on the most recent definitions of dystonia and chorea/athetosis(15)

1.4 MEASUREMENT PROPERTIES

Clinicians and researchers use measurement to understand and evaluate different characteristics in individuals. Measurements are only useful if the outcome data is accurate, consistent, and meaningful. Reliability and validity are measurement properties describing to what extent a test instrument produces data that meet those criteria. Choosing a suitable test instrument is not always straightforward, since reliability and validity are dependent on the population of interest as well as the situation. Re-evaluation of reliability and validity is necessary when using a test for a new target population; for example in children instead of adults, or when only parts of the test are used, or if the test is slightly changed.

1.4.1 Reliability

The first quality criteria for a test instrument is that it produces consistent results, meaning that if the test is administered at a later occasion to the same test takers, the outcome will be similar. Reliability refers to the consistency in the measurement; to the precision of the test result and to the degree the test results are free from random error, and thus can be replicated(92, 93). The consistency may be between two different raters (inter-rater reliability), or the same rater on different occasions (intra-rater reliability), or between two repeated assessments over a short period of time, when no change in the measured variable is expected (test–retest reliability). Internal consistency reflects the extent to which different items in the test instrument measure various aspects of the variable of interest, and nothing else.

1.4.1.1 Reliability coefficients for test–retest reliability and rater reliability

The intraclass correlation coefficient (ICC) is a preferred index for test–retest, inter-rater and intra-rater reliability for quantitative data, since it reflects both agreement and correlation(92, 94). The ICC describes how strongly data in a specific group resemble one another; i.e., scores from the same group tend to be similar. The ICC value ranges from –1 to 1, where 0 indicates no correlation (random) and 1 indicates perfect correlation. Negative values exist, and would be interpreted as having less correlation than would be random. When interpreting ICC values, you must be cautious. In the case of a high value, it is important to check that an adequate model was chosen (see below). In the case of a low ICC value, it may be important to check the variability among subjects scores. A homogenous sample, with a low variability of scores, will result in low ICC values. Low ICC values may also be explained by a low number of study subjects or raters(94).

There are several different ICC models(94). To choose the right one is somewhat complicated, and there are many aspects to address. Are all subjects/participants scored by the same raters? Are the raters a specific sample of raters, or are they randomized? In the intended measurement protocol, is the outcome based on a single rater or is it based on a mean of multiple raters? And lastly, do you consider absolute agreement or consistency to be more important? Based on the answers of these questions, you may choose between three different models (1-way random-effects model, 2-way random-effects model, and the 2-way mixed-effects model), and for each model one of two types (single rater or multiple raters) and one of two definitions (consistency or agreement)(92).

1.4.2 Validity

The second quality criteria for a test instrument is measurement validity. Validity concerns the extent to which a test measures what it is intended to measure(95, 96). Here, the focus is the objective or the aim of the test instrument, and the ability to make inference from the test scores. There are different variants and definitions of validity, depending on how the test instrument is supposed to be used and if there are any existing instruments to compare it with. Traditionally, four main variants of validity are considered to be useful; face validity, content validity, criterion-related validity, and construct validity(95, 96).

Face validity is the weakest form of validity, and indicates that a test appears to test what it is supposed to. *Content validity* refers to what extent the content of the test instrument represents the trait to be measured adequately. Does the test include all relevant aspects of the characteristic to be measured? This is best evaluated with experts in the field. *Criterion-related validity* indicates that the outcome of the test can be a substitute for a more comprehensive/standard test. Concurrent validity is one type of criterion-related validity, and can be evaluated if there is a validated ‘gold standard’ as a comparison. The same trait is assessed with the new test and the ‘gold standard’ simultaneously, and the correlation is calculated.

Construct validity is a measure of the ability of the test to measure an abstract construct. Part of construct validity is based on the content validity, but beyond that, construct validity depends on the underlying theoretical context. Construct validity provides evidence to support (or

refute) the theory behind the construct under investigation. Construct validation of a questionnaire or an assessment scale is an ongoing process, and there are different methods of construct validation. Known groups' validation may show that the scale is able to differentiate between individuals with more or less of the trait. Convergent and discriminant validity are other aspects of construct validity, when the outcome measure of the test is compared to a test with a similar trait (convergent) or a different trait (discriminant)(95, 96).

1.4.3 Theories of measurement

When measuring a human characteristic, a test instrument with a combination of different test items, covering different aspects of the trait of interest, is often used. For example, a test instrument to assess dystonia is based on a combination of test items for different body regions, in action as well as during rest. The scores from all these different test items are summed up to a total score—an outcome measure—which is supposed to be a measure of the underlying trait, in this case the dystonia.

1.4.3.1 Classical test theory

Classical test theory (CTT) is a traditional approach to evaluate reliability and validity of a test instrument based on several items(97). CTT assumes that that for each test, or each test item, the observed score is a combination of an underlying true score (on the trait of interest) and a random error. The random error is not related to the magnitude of the true score, and random errors will cancel each other out (approach zero) with repeated measurements. If these assumptions are correct, the average score is a good estimate of the true score if enough measurements are taken. Furthermore, the observed score of the test instrument approaches the true score when the number of test items increases. Accordingly, the reliability of the test instrument increases when the number of test items increases.

There are some limitations with CTT, however(97). First, the reliability and validity found for a test instrument only applies to the specific group of people who took the test (97). The psychometric properties will automatically change if the test is applied to participants of another population. Similarly, the psychometric properties will change if the test items are slightly changed or if some test items are deleted.

Another limitation of CTT is the assumption of item equivalence, where each test item contributes equally to the total score of the test instrument, although different test items may correlate differently with the trait of interest(97).

Furthermore, in CTT the standard error of measurement (SEM) is assumed to be the same for all individuals, all over the scale, and hence independent of the amount of the trait of interest. Statistics and mathematical models, however, show us that this is incorrect. SEM increases with higher and lower total scores and is smallest in the middle(97). Subsequently, total scores and SEM from high or low performers must be interpreted with caution. And last, the outcome measure is ordinal data, making comparison between individuals complicated.

1.4.3.2 Item response theory and Rasch measurement model analysis

To overcome the limitations connected with CTT, two independent research groups in the 1960s used another approach to evaluate the measurement properties of a test instrument, and subsequently, item response theory (IRT) and Rasch measurement model analysis were developed(97). These models focus on the information from the different test items, whereas CTT focuses on the total score of the test instrument. Item response theory (IRT) and Rasch measurement model analysis were developed separately but share important features. Some researchers consider Rasch analysis one IRT model, whereas representatives of the Rasch analysis model emphasize their (fundamental) differences(98).

Rasch analysis is based on the relationship between an individual's ability and the degree of difficulty of the test item(99). This relationship can be described as a probability of a specific response to a test item, such as probability of success. The person (ability) and the item (difficulty) can be ordered on the same scale, according to probabilities of success. The aim is to measure the ability of the person, by estimating the location of the individual on this scale based on the responses of the test items. Rasch analysis has two hard assumptions; 1) that the scale only measures one trait (unidimensionality), and 2) that the outcome of each test item is independent of the others (local independence). If these assumptions are not fulfilled, it is impossible to predict the probability of success, and the items and persons cannot be ordered along the same scale.

Rasch analysis can be used by test developers to examine how well test items (and persons) 'fit' the Rasch model. Data from individuals having undergone a new test instrument are analyzed with the Rasch measurement model analysis. If the data fit the model, this is evidence of unidimensionality and local independence of the test instrument. Fit statistics and principal component analysis are used in this process and will be described more in detail in the methods section of this thesis. Visual inspection of item characteristic curves may further guide the test developer to improving the test instrument. The item characteristic curve is visualized in a graph as an S-shaped curve, where the x-axis is represented by a person's ability and the y-axis by probability of success. Redundant or non-fitting test items may be deleted from the test instrument.

There are several advantages of a test instrument based on a Rasch measurement model analysis. First, the outcome measure is interval level data. Each item and each person will receive a measure on a logit scale, based on probability of success. One logit is the distance that increases the odds of success by a factor of 2.718 (the base of natural logarithms). Additionally, each item has a unique contribution to the total score of the test instrument; i.e., some items have a greater impact on the total score than others. Furthermore, each item has a unique SEM. Another advantage is that for high or low performers, you may choose to delete test items from the test in order to prioritize those items that offer the most information about the individual.

2 AIMS OF THE THESIS

The overall aim of this thesis was to enable and facilitate assessment of dystonia and choreoathetosis in children and young adults, and to elucidate some clinical aspects in two rare pediatric-onset dystonia disorders.

The specific aims of the studies were as follows:

Study I: To examine the natural long-term development of BPTI, its relationship to migraine and other paroxysmal disorders, and the frequency of mutations in four candidate genes (*CACNA1A*, *PRRT2*, *ATP1A2*, and *SCN1A*).

Study II: To examine the clinical outcome after pallidal DBS in DYT-THAP1 dystonia and to examine factors that may predict the amount of subsequent dystonia reduction.

Study III: To examine the inter-rater and test–retest reliability and concurrent validity of the DIS in children and young adults with inherited or idiopathic dystonia.

Study IV: To develop a shortened, user-friendly version of the DIS, the DIS-II, and to evaluate its construct validity and reliability with a Rasch measurement model analysis. A secondary aim was to change the relative outcome measure in the original DIS to an absolute outcome measure in DIS-II.

3 METHODS

3.1 STUDY DESIGNS

Studies I and II were case series of individuals with two rare pediatric-onset movement disorders, with a long-term follow-up. Study I was a prospective longitudinal clinical study with two on-site follow-ups, whereas Study II was a retrospective, multi-center study evaluating individual data obtained from patient medical records. Studies III and IV were cross-sectional, two-center studies, evaluating (and developing) an assessment scale. In Study III, inter-rater reliability, test–retest reliability, and concurrent validity of an existing assessment scale were explored for a new population. In Study IV, a shortened version of that assessment scale was developed, and the Rasch measurement model analysis was used to explore rating scale functioning, internal scale validity, targeting, and internal consistency.

3.2 PARTICIPANTS

Swedish participants had given informed written and/or oral consent, either themselves or by their caregiver. For participants outside Sweden, legislation regarding consent in the relevant country was followed. All studies were approved by the appropriate Ethical Review Board. An overview of the participants is given in Table 1.

Table 1. Overview of participants in Studies I–IV

Study	Total number of participants, N	Inclusion criteria	Exclusion criteria	Female sex, N (%)	Participants recruited from country (N)
I	12	- diagnosis of BPTI at Astrid Lindgren Children's Hospital between 1998-2005	none	8(67)	Sweden (12)
II	14	- dystonia *and* - mutations/sequence variants in *THAP1* *and* - GPi-DBS with ≥6 months follow-up at inclusion	none	9(64)	England (2) France (3) Germany (1) Italy (5) Sweden (3)
III	20	- inherited or idiopathic dystonia *and* - 5–25 years of age	- paroxysmal disorders - dopa-responsive dystonias	9(45)	Belgium (6) Sweden (14)
IV	123	- inherited/acquired/idiopathic dystonia or dyskinetic cerebral palsy *and* - 5–25 years of age *and* - videotaped according to the DIS	none	49(40)	Belgium (100) Sweden (23)

N =number, BPTI =benign paroxysmal torticollis of infancy, GPi-DBS =pallidal deep brain stimulation

3.2.1 Recruitment

In Study I, all individuals having received a diagnosis of torticollis at Astrid Lindgren Children's Hospital between 1998 and 2005 were identified. Medical records were reviewed in 2005 by a pediatric neurologist and a pediatric physiotherapist, to delineate those who had received the diagnosis of benign paroxysmal torticollis of infancy (BPTI) by their treating pediatrician. Thirteen children met the inclusion criteria, and 12 agreed to participate in the study.

Participants in Studies II–IV were convenience samples. Two of the Swedish participants in Study II participated in Study III, and all participants in Study III participated in Study IV.

In Study II, individuals with DYT-THAP1 dystonia, who had received treatment with pallidal deep brain stimulation were recruited from 5 European centers (Stockholm, Sweden; Milan, Italy; Montpellier, France; London, England and Cologne, Germany). They were identified by pediatric neurologists or neurologists clinically experienced in movement disorders through personal knowledge.

In Study III, individuals 5–25 years of age, with a diagnosis of inherited or idiopathic dystonia, were recruited from Stockholm, Sweden and from Leuven, Belgium. They were identified by pediatric neurologists through personal knowledge.

In Study IV, individuals 5–25 years of age, with a diagnosis of dyskinetic cerebral palsy or inherited/acquired/idiopathic dystonia who had been recorded according to the DIS film protocol in Leuven, Belgium or Stockholm, Sweden were recruited. The participants were identified by pediatric neurologists through personal knowledge.

3.2.2 Sample size

Study I and II were case series, where the aim was to include as many participants as possible. In Study I, all individuals who received the diagnosis BPTI at Astrid Lindgren Children's Hospital during the years 1998–2005 were invited to participate. Thirteen children met the inclusion criteria for the study, but one boy declined to participate. Eight girls and 4 boys were included in the study, but one boy was lost during the second follow-up. In Study II, 14 individuals were recruited through convenience sampling, nine females and five males: Stockholm n =3, Milan n =5, Montpellier n =3, London n =2, and Cologne n =1.

In Study III, the initial goal was to recruit 25 participants, since that was in line with the number of participants in the original study of the DIS assessment scale. Inherited and idiopathic dystonia are unusual conditions, and during the study we realized that cost–benefit was in favor of fewer participants. Calculating sample sizes for reliability studies based on ICC values is not straightforward, but some researchers have developed tables to simplify the calculation(100, 101). To use these tables, you need to specify the significance level α and power β (often 0.05 and 0.80), and then decide how many observations per study participant you are going to perform, a minimum level of acceptable ICC-value (null hypothesis) and the ICC-value you aim to receive as a result (alternative hypothesis)(100, 101). According to the methods of Bujang et al., we calculated the required sample size to be at least 18 individuals, with the null hypothesis being that the ICC values would be >0.7 and the alternative hypothesis that the ICC

values would be >0.9, alpha being 0.05, with power set to 90% with three observations per study participant(100). Consequently, we stopped including more individuals when 20 participants, 11 males and nine females, (Belgium n =6, Sweden n =14), had been assessed.

In Study IV, a sample size of 100-150 persons should indicate item calibration within ±0.5 logit with 95% confidence according to Linacre(102). A total of 123 participants from Belgium (n =100) and Sweden (n =23); 74 males and 49 females were included in the study.

3.3 PROCEDURE AND DATA COLLECTION

3.3.1 Study I

Study I was a prospective longitudinal follow-up case series of children with benign paroxysmal torticollis of infancy (BPTI). At a first follow-up, during the years 2005–2007, 12 children—who had been diagnosed with BPTI in 1998–2005 at the neuropediatric unit at Astrid Lindgren`s Children hospital—were examined and evaluated by a pediatric neurologist and a pediatric physiotherapist. Children ≤3 years of age were evaluated with the Peabody Developmental Motor Scale-2 and children ≥4 years of age were evaluated with the Movement Assessment Battery for Children. Information regarding the torticollis episodes was gathered during those visits.

In a second follow-up, during the years 2014–2015, 11 children and their caregivers were interviewed by a pediatric neurologist (the author). One boy was lost to the second follow-up. The focus of the interviews was the development of the child, the possible occurrence of migraine or other paroxysmal disorders in the child, and any family history of migraine or other paroxysmal disorders over three generations. Blood or saliva samples were gathered from the children for genetic analysis of candidate genes. In 2018, after the genetic analysis was finished, all data were analyzed, and the manuscript was written.

3.3.2 Study II

Study II was a retrospective follow-up case series, a multi-center study of individuals with DYT-THAP1 dystonia treated with pallidal deep brain stimulation (DBS). The research team in Stockholm invited clinicians (neurologists/neurosurgeons) from London, Montpellier, Cologne, and Milan to include patients with DYT-THAP1 dystonia treated with pallidal DBS for at least six months in this study. These clinicians were free to invite other specialists that could contribute to the study when appropriate (geneticists, neurologists, pediatric neurologists).

A data base template, with a focus on demographic, genetic, and pre-and postoperative data, was created in Stockholm by two pediatric neurologists (the author of this thesis and her principal supervisor) and a neurosurgeon and distributed to the participating centra. Data were collected retrospectively from medical records by the clinicians, according to the data base template. In 2017–2018, data collection was preformed, and in 2018–2019, data were analyzed, and the manuscript written.

3.3.3 Study III

Study III was an instrument evaluation; a cross-sectional, two-center study of the Dyskinesia Impairment Scale (DIS) in inherited or idiopathic dystonia. In 2016–2018, individuals with inherited or idiopathic dystonia from Stockholm, Sweden (n =14) and Leuven, Belgium (n =6), were invited to two visits, with a maximum of two weeks in between.

At the first visit, participants from Stockholm, Sweden, were interviewed, examined, and filmed according to a standard DIS video protocol, by a pediatric neurologist (the author) and a pediatric physiotherapist at the motion analysis laboratory in Astrid Lindgren Children's Hospital. The interview aimed to obtain demographic information regarding age at onset, functional gross motor and manual ability, and to complete the BFM assessment. The examination aimed to evaluate the joint range of motion and possible occurrence of spasticity. Genetic information was gathered from medical records. In Leuven, Belgium, participants were examined and filmed by a movement scientist in their habitual environment (school, home), and demographic/genetic information was gathered from their medical records.

At the second visit, participants were filmed according to the same video protocol again. The research team perceived one girl to be uncomfortable during the first visit and she was therefore excluded from the second filming.

In 2019, during a 4-week period, the videos were scored by three raters: one pediatric neurologist (the author), one pediatric physiotherapist, and one movement scientist. The raters had attended a DIS instructional course in Belgium, conducted by the constructor of the DIS, and all had clinical experience in dystonia. Before scoring the participants in the study, the three raters scored four training videos together. The raters independently scored all videos from visit 1 according to the DIS and the BFM-M. The videos from visit 2, were randomly distributed among the raters, and each video was scored by one rater according to the DIS.

The scores from visit 1 were used to evaluate inter-rater reliability (each participant was scored by all three raters with DIS) and concurrent validity (correlation between DIS-D scores and BFM-M scores from each rater). The DIS score from visit 2 was paired up with the same rater's DIS score from visit 1 to evaluate test–retest reliability. In 2020, the data were analyzed, and the manuscript written.

3.3.4 Study IV

In Study IV, a cross-sectional, two-center study, the DIS-II was developed and evaluated regarding construct validity and reliability. The study consisted of three phases: i) development of the DIS-II; ii) data collection for a Rasch measurement model analysis; and iii) evaluation of construct validity and reliability of the DIS-II using Rasch measurement model analysis.

i) Development of DIS-II: The research team invited clinicians and researchers familiar with the DIS to attend an advisory expert meeting, to elucidate strengths and weaknesses of the DIS. This online, whole-day meeting, with 21 experts from Belgium, the Netherlands, Germany, Turkey, and Sweden, took place in February 2021. A poll was prepared by the research team, with questions covering different aspects of the DIS, such as body regions included, actions/rest postures, and the rating scale construct. The attendees answered the poll anonymously and

afterwards there was an open discussion. A preliminary DIS-II was developed by the research team (IV, AD, KT, CL, LKS, EM) during spring 2021. Results from the expert meeting, previous existing psychometric properties regarding the original DIS, clinical aspects, utility, and scoring feasibility, were considered in the process.

ii) Data collection for the Rasch measurement model analysis: data were gathered from different sources; in total 123 unique individuals with dystonia, 5–25 years of age, from Leuven, Belgium (n =100) and Stockholm, Sweden (n =23) were included. For 97 participants, DIS scores from earlier studies could be reused (whereof 20 of those were included in Study III). For 26 participants, DIS videos from a clinical setting were available (17 from Leuven and 9 from Stockholm), and were scored during December 2020–January 2021 by the research team as a part of this study.

iii) Evaluation of construct validity and reliability of the DIS-II, using Rasch measurement model analysis: the research team had several physical and online meetings from June 2021–December 2021 where Rasch measurement model analysis was performed on the preliminary DIS-II. Measurement properties and clinically meaningful aspects were considered, and a final DIS-II was designed.

3.4 GENETIC ANALYSIS

In Study 1, a genetic analysis was performed. Whole-exome sequencing was performed on genomic DNA, extracted from blood or saliva. Sequence variants detected in *CACNA1A*, *PRRT2*, *ATP1A2*, and *SCN1A* were verified using Sanger sequencing.

3.4.1 Sanger sequencing

Sanger sequencing has been the most common sequencing method in the last 40 years, and today it is a gold standard method, often used to validate the results found with next generation sequencing(103, 104). Sanger sequencing involves electrophoresis and is based on random incorporation of modified di-deoxynucleotides that terminates in vitro DNA replication. The method requires single-stranded DNA (the DNA of interest), a DNA primer, a DNA polymerase, normal deoxynucleotide triphosphates (dNTP; adenine guanine, cytosine, and thymine), and modified di-deoxynucleotide triphosphates (ddNTP). The modified ddNTP are fluorescently labelled, which enables identification in sequencing machines.

The first step involves denaturation of the double-stranded DNA into two single-stranded DNA molecules. Then, a DNA primer is used as a starting point for the DNA sequence. The DNA polymerase adds normal dNTPs to the growing strand, and the process stops when a modified fluorescent ddNTP is randomly incorporated into the DNA. The resulting DNA fragments, of different lengths, are separated using electrophoresis, and the relative position of the fluorescent ddNTPs are used to read the DNA sequence.

3.4.2 Whole-exome sequencing

Whole-exome sequencing (WES) is a next-generation sequencing method that involves sequencing of the protein-coding regions of the genome(105, 106). The exome represents about 1%–2% of the total human genome but contains about 90% of known disease-related variants. The critical difference between Sanger sequencing and WES is the sequencing volume. The Sanger only produces one single DNA fragment at a time, whereas WES is massively parallel, sequencing millions of fragments simultaneously.

The technology used in the thesis, was sequencing by synthesis from Illumina (Illumina Inc, San Diego, CA, USA). The process for detecting different types of variants/mutations consists of four steps: 1) library preparation, 2) cluster generation, 3) sequencing, and 4) data processing/analyzing including comparison to a reference sequence. The first step, library preparation, includes fragmentation of the DNA and then 'capturing' of all protein coding exons in the genome. The subsequent cluster generation is a process where the DNA fragments are bound to a surface and PCR amplified to create a cluster of identical DNA fragments. In the next step, sequencing is performed with the clusters of DNA fragments as templates. The DNA polymerase adds fluorescent nucleotides one by one to the growing DNA strand. In the last step, the data generated from the sequencing is processed and compared with a reference genome, to detect differences between the two.

3.5 ASSESSMENT SCALES

3.5.1 The Peabody Developmental Motor Scale-2

This test was used by a pediatric physiotherapist in the first follow-up in Study I, to assess fine and gross motor function in children <4 years of age. It is a comprehensive, norm-referenced test for children from birth to five years of age, and consists of six subtests: reflexes, stationary gross motor skills, locomotion, object manipulation, grasping, and visual motor integration(107).

3.5.2 The Movement Assessment Battery for Children

This test was used by a pediatric physiotherapist in the first follow-up in Study I, to assess fine and gross motor function in children ≥4 years of age. This assessment consists of eight test items, divided into three performance areas: manipulative skills, ball skills, and balance skills(108).

3.5.3 The Dyskinesia Impairment Scale (DIS)

The DIS is a measurement tool used to evaluate dystonia (DIS-D) and choreoathetosis (DIS-CA) in children and young adults with dyskinetic cerebral palsy, Figure 1(88). The patient is filmed, according to a standard video protocol, with a total of 26 video sequences of about 30 seconds each. Twelve body regions are assessed during one rest posture and two actions each,

Table 2. For each rest posture and each action, the duration (amount of time) and amplitude (range of motion) of dystonia/choreoathetosis are assessed on a 5-step rating scale, 0–4. A score of 0 indicates no dystonia/choreoathetosis and a score of 4 indicates that the dystonia/choreoathetosis is present ≥90% of the time/of the range of motion. The sum of raw scores range between 0–576 for the DIS, which is based on the sum of raw scores for the two subscales DIS-D (range 0–288) and DIS-CA (range 0–288). Actions that the individual can perform are included in the assessment, others are removed. Outcome measures are presented as raw scores, or as percentage scores, based on those rest postures/actions that were performed.

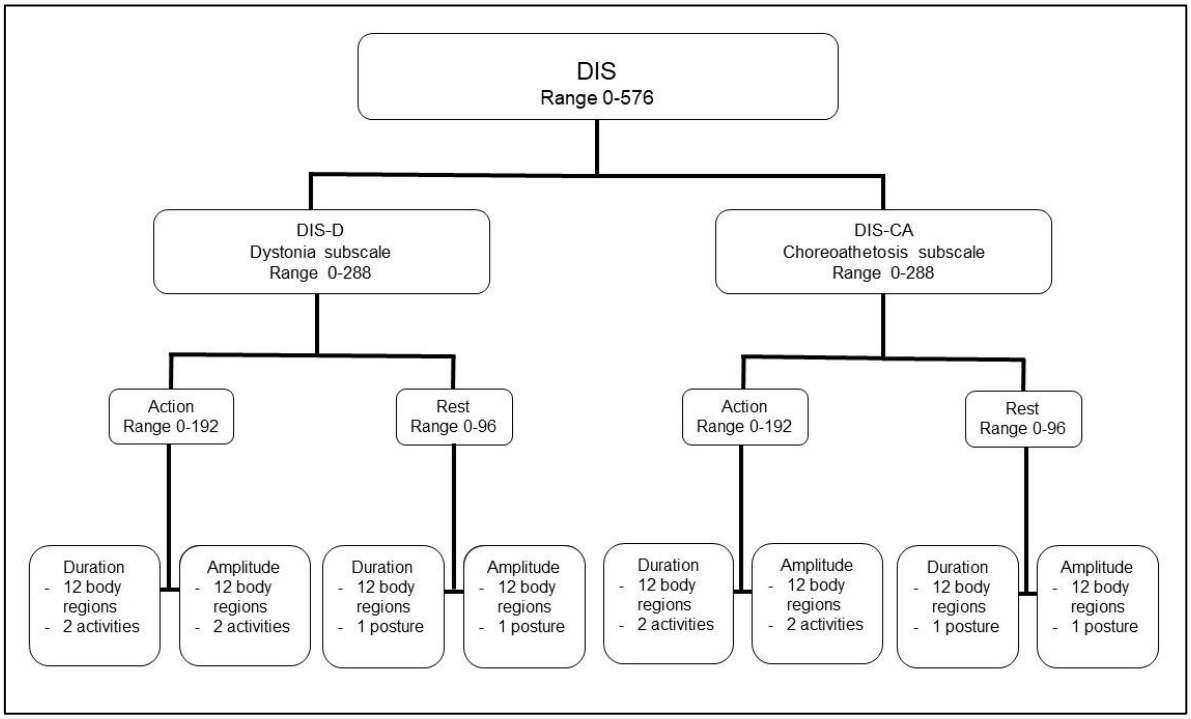


Figure 1. The Dyskinesia Impairment Scale: DIS =Dyskinesia Impairment Scale, DIS-D =dystonia subscale, DIS-CA = choreoathetosis subscale. Ref J.Clin.Med. 2020, 9, 2597; doi:10.3390/jcm9082597

Table 2. Overview of body regions, actions, and rest postures in the Dyskinesia Impairment Scale.

Body region	**Action**	**Rest posture**
Eye	Eye tracking Eye blinking	Sitting at rest, close-up view
Mouth	Open mouth Speech	Sitting at rest, close-up view
Neck	Rotation Latero-flexion	Sitting at rest, close-up view
Trunk	Bending trunk Active sitting	Sitting at rest, frontal view
Proximal arm (left/right)	Abduction Reaching a pen (supine position)	Sitting at rest, frontal view
Distal arm (left/right)	Grasping and moving a pen (sitting position) Grasping and moving a cup (sitting position)	Sitting at rest, frontal view
Proximal leg (left/right)	Standing Rolling	Lying in rest
Distal leg (left/right)	Heel-toe tap Rolling	Lying in rest

3.5.4 The Burke–Fahn–Marsden Dystonia Rating Scale (BFM)

The BFM is an assessment scale developed for the evaluation of dystonia in adults with isolated dystonia (primary dystonia), and consists of two subscales, the movement scale (BFM-M) and the disability scale (BFM-D)(86). The BFM-M is based on filming of the patient, whereas the BFM-D is based on patients' or caregivers' report of disability in activities in daily living.

The BFM-M evaluates dystonia in nine body regions, including the eyes, mouth, speech/swallowing, neck, trunk, and right and left sides of the upper and lower extremities. The severity factor (scored 0–4) refers to the amount of dystonia that is present in each body area, and the provoking factor (scored 0–4) refers to which situation elicits the dystonia (rest/action). The severity factor, the provoking factor, and a weighing factor (0.5 or 1) for each body area are multiplied, and then the total sum is calculated. The weighing factor 0.5 is used for the eyes, the mouth, and the neck to down-weigh them. The total score of the BFM-M ranges between 0–120, where a low score indicates less dystonia.

The BFM-D evaluates the impact the dystonia has on speech, handwriting, feeding, eating/swallowing, hygiene, dressing, and walking. Each area assessed, is given a score between 0–4, except walking which is scored 0–6, subsequently leading to a score range of 0–30, where a low score indicates less dystonia.

3.6 STATISTICAL METHODS

An overview of statistical methods is presented in Table 3.

Table 3. Overview of statistical methods in Studies I–IV.

Study	Aim	Design	Data analysis
I	Longitudinal development	Prospective longitudinal follow-up case series	Descriptive statistics (mean, standard deviation, median, range)
II	Differences between paired observations, correlation of variables	Retrospective follow-up case series, multi-center study	Descriptive statistics, Wilcoxon signed rank test, effect size, Spearman correlation coefficient
III	Inter-rater and test–retest reliability, concurrent validity	Instrument evaluation, cross-sectional, two-center study	Descriptive statistics, intraclass correlation coefficient, standard error of measurement, minimal detectable difference, Spearman correlation coefficient
IV	Construct validity (internal scale validity) and reliability (internal consistency)	Instrument evaluation, cross-sectional, two-center study	Descriptive statistics, Rasch measurement model analysis (goodness-of-fit statistics, principal component analysis, rating scale functioning, targeting, and reliability)

For Studies I–III, the statistical analysis was performed by using IBM SPSS software 25. For Study IV, a Rasch measurement model analysis was made, using WINSTEPS software version 4.8.2.0 (John M Linacre, 2015).

3.6.1 Descriptive statistics

Descriptive statistics are used to characterize a given set of data, regarding central tendencies (measure of the center) and variability (measure of the spread)(109). Depending on the type of data, different measures are appropriate. For ordinal data, medians and interquartile ranges (IQR) are used, whereas for interval data, means and standard deviations (SD) are most often the right choice. In some cases, however—for example skewed interval data—medians and IQR is a better option, since the median value is less affected by outliers. Descriptive statistics were used in Studies I–IV.

3.6.2 Comparing groups and examining correlations

3.6.2.1 Wilcoxon signed rank test

Wilcoxon signed rank test is a non-parametric statistical hypothesis test that examines the difference between paired observations(110). The Wilcoxon signed rank test can be compared with the paired t-test and can be used when the assumptions of that test are not fulfilled. The Wilcoxon signed rank test is recommended when the data level is categorical (nominal or ordinal data) or when the population data (the differences between the pairs of data) does not have a normal distribution(110). The test examines whether the median of the differences between paired observations is zero in the population from which the sample is drawn. In Study II, the Wilcoxon signed rank test was used for ordinal data to analyze potential differences before and after the intervention.

3.6.2.2 Effect size

Effect size helps us to understand the magnitude of a found difference, whereas statistical significance examines whether a finding is likely to be due to chance(111, 112). In Study II, effect size was calculated using the formula $r = z/\sqrt{n}$, where z =the result obtained from the Wilcoxon signed rank test, and n =number of observations(113). An effect size of r =0.1 was considered small, r =0.3 medium, and $r \geq 0.5$ large.

3.6.2.3 The Spearman correlation coefficient

The Spearman correlation coefficient is used instead of the Pearson correlation coefficient when correlations are based on ranks of data(114). The assumption of a normal distribution is not required. In Studies II and III, the Spearman correlation coefficient was used to explore correlations between different variables, and the size of the correlation coefficient was interpreted as negligible (0.0 <0.3), low (0.3 <0.5), moderate (0.5 <0.7) or high (0.7–0.9).

3.6.3 Measures of reliability

3.6.3.1 The intraclass correlation coefficient

The intraclass correlation coefficient (ICC) is a preferred reliability coefficient for inter-rater reliability and test–retest reliability for quantitative data, since it reflects both agreement and correlation(115). The ICC describes how strongly data in a specific group resemble one another: i.e. scores from the same group tend to be similar.

In Study III, the reliability coefficient ICC was used to assess inter-rater and test–retest reliability. An ICC value of 0 indicates no correlation (random) and 1 indicates perfect correlation. The quality criteria we used in Study III were; ICC values >0.90 were interpreted as excellent, values between 0.75 and 0.90 as good, values between 0.60 and 0.75 as moderate and values <0.60 as poor(115). For more information about interpretation of ICC values and different ICC-models, see 1.4.1.

Inter-rater reliability: The 2-way random-effects model was used to examine inter-rater reliability. That model was chosen since all raters scored all persons, and since generalizing of the reliability result to other raters with the same experience was planned(115). Furthermore, the type "single rater" was chosen, since the DIS when used in clinic as well as in research is evaluated with a single rater. And lastly, absolute agreement was chosen, since that is of great importance in an assessment scale.

Test–retest reliability: The 2-way mixed-effects model was used to examine test–retest reliability, since the only rater of interest here, is the one doing the two evaluations, with a short period of time inbetween, where no clinical change is expected(115). Single rater and absolute agreement were kept for the same reasons as in inter-rater reliability.

3.6.3.2 The standard error of measurement

The standard error of measurement (SEM)—the standard deviation of the measurement error—is a reliability measure(116, 117). The SEM describes how accurate a test score is, where a low SEM indicates high score accuracy, and a high SEM indicates low score accuracy.

The SEM may be derived from the standard deviation of the observed scores and the reliability coefficient, using the formulae $SEM = SD * \sqrt{(1–ICC)}$(116, 117). An advantage of the SEM, compared to the dimensionless reliability coefficient ICC, is that it is expressed in the same unit as the test score, and is therefore easier to interpret. From the SEM, a confidence interval around a test score can be calculated. As an example, a 95% confidence interval would be [observed test score] +/– 1.96 * SEM.

In Study III, the SEM for test–retest was calculated for the DIS, the DIS-D, and the DIS-CA. The standard deviation was calculated from the scores in the first of the two assessments for each scale. The ICC value was derived from the test–retest for each scale.

3.6.3.3 Minimal detectable difference

Minimal detectable difference (MDD) is the smallest difference that can be statistically detected in a study, is always bigger than the SEM, and is dependent on your choice of significance level(118). In Study III, MDD was calculated using the formulae MDD =SEM * 1.96 * √2(118).

3.6.4 Measures of validity

In Study III, concurrent validity was examined(95). The result from the DIS-D was compared to a gold standard, the BFM-M, and the Spearman correlation coefficient was calculated, since the data from the two scales are ordinal(114).

3.6.5 Rasch measurement model analysis

Rasch analysis is a model that can be used by test developers to design a new test instrument or evaluate an existing one(99, 119). The fundamental concept of this model is that an interval level measurement scale can be constructed out of the relationship between two ordinal level datasets (person ability and item difficulty). This is a great advantage for a test instrument, since data on interval levels facilitate comparison, when all scale steps are equal, and persons and/or items can be compared at all points of the scale.

The Rasch model is a mathematical model that can be used to analyze a dataset with responses to a test instrument. Response patterns are analyzed and the probability of success is determined for each person and each test item. The probability of success depends on the ability of the person and the difficulty of the item. Two important assumptions behind the model are: 1) all persons are more likely to succeed on an easy item compared to a more difficult item, and 2) all items are more likely to be scored correctly by more able persons than less able persons(99).

The model is dependent on unidimensionality of the test instrument, meaning that the test instrument only assesses one characteristic (but assesses all aspects of that characteristic). Another important requirement is the local independence of test items, where a correct or wrong reply to one test item should not lead to a correct or wrong reply to another item.

The Rasch measurement model analysis is performed on a dataset containing responses from different individuals to a test consisting of several test items. Each person and each test item receives a ‘measure’ on the logit scale, determined by the probability of success, and can be placed along an ideal straight line according to difficulty/ability. This straight line is, of course, theoretical and observed response patterns deviate from this ideal line. Rasch measurement model analysis investigates how well the persons and items match this straight line, and misfitting items/persons can be identified, and changed or removed to improve the test.

A bubble chart of item difficulty and personal ability versus their fit/misfit values may be helpful in identifying troublesome items/persons (Figure 2). Each bubble represents an item or a person, and their placement along the y-axis is based on the difficulty of the item and the ability of the person. Items located at the bottom are easier, whereas those located at the top are more difficult. Persons located at the bottom are less able, and those located at the top are more able. A person with the same measure as an item; i.e., with the same location along the y-axis, has 50% probability of scoring the item correctly. The ordering of the item can be interpreted as a developmental pathway. If the ordering of items is logical according to the knowledge of the underlying characteristic we are measuring (the latent trait), that is evidence supporting construct validity of the test instrument. Furthermore, if the data is considered to fit the model well enough, and accordingly the persons and items match the ideal line well enough, the Rasch assumptions are fulfilled, and the test instrument is unidimensional.

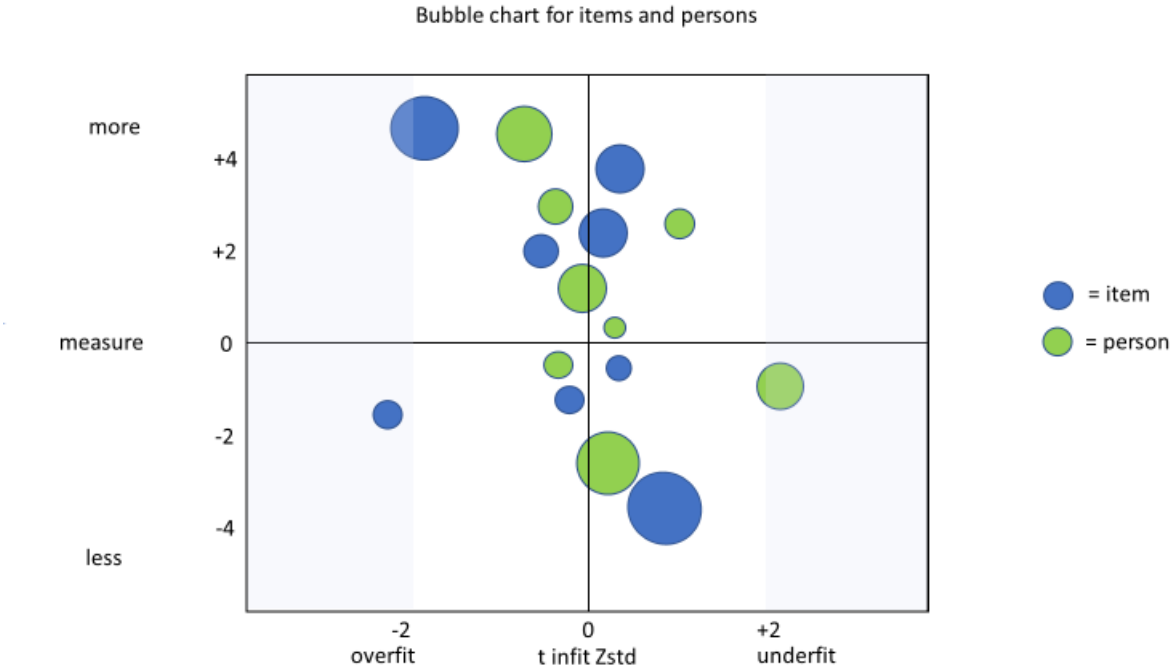


Figure 2. Bubble chart for items and persons; A bubble chart is an illustration of the developmental pathway along the latent trait and shows measures and fit statistics of items and persons. Each bubble represents an item or a person, and the size is proportional to the standard error of the item difficulty/person ability. Zstd =standardized as a z-score

In Study IV, the Rasch measurement model analysis was used to examine construct validity and reliability of the DIS-II and the two subscales DIS-II-D and DIS-II-CA. The latent trait is the dyskinesia; a high person measure indicates more dyskinesia, and a high item measure indicates less dyskinesia (the item is more 'difficult', and elicits dyskinesia less easily). Specifically, aspects of rating scale functioning, unidimensionality, targeting and internal consistency were examined. Furthermore, the Rasch analysis was used to convert raw scores from the DIS-II, DIS-II-D, and DIS-II-CA to logit scores.

3.6.5.1 Choice of model

In Rasch analysis, there are different mathematical models depending on the response structure of the test items. The dichotomous model is being used when the response options are, for example, 'right' or 'wrong'. In our study, a polytomous model was appropriate since the test items in DIS-II are scored on a rating scale with 3 or 4 different options. The polytomous Andrich Rating Scale model can be used when all items share the same category rating scale structure. In Study IV, however, the grouped rating scale model was chosen, since rest items are one group of items with a scoring scale between 0–2, and action items are another group of items scored between 0–3(120).

3.6.5.2 Rating scale functioning

Quality criteria used for a well-functioning rating scale in Study IV were (I) average measure advancing monotonically by category, (II) item step calibration increasing with higher categories, (III) each category being represented by a curve with a distinct peak upon visual inspection, (IV) frequency for all four categories exceeding 10 ratings and (V), the outfit mean-squares should not exceed 2.0 for any category(121, 122).

3.6.5.3 Unidimensionality

Point-measure correlations are usually examined first when investigating unidimensionality. Item point-measure correlations are the correlations between the observed scores for an item and the person measures, and thereby a measure of how well the item measures represent the latent trait. In Study IV, positive item point-measure correlations were considered sufficient(123). If all correlations are positive, fit statistics can be used for more thorough analysis.

Fit statistics describes the extent to which responses adhere to the Rasch model and are based on residuals. Residuals represent the difference between the observed and the expected value according to the Rasch model(124). For each item and person, the sum of all squared residuals is divided by the number of persons/items contributing to the data. The result is reported in an unstandardized and a standardized form, as mean squares (MnSq) and z-values respectively(124). The MnSq are modelled by the Rasch algorithm to have a mean of 1, and the expected MnSq therefore is equal to 1. In the standardization of the MnSq, they are transformed to a z-distribution, with a mean of 0, and a standard deviation of 1.

In Study IV, fit statistics were evaluated by examining MnSq and z-values. Fit statistics are differentiated into infit- and outfit statistics. Infit statistics are information-weighted and sensitive to responses with difficulty close to the ability of the person, whereas outfit statistics are sensitive to responses far from ability level(124). As misfitting infit statistics are a greater threat to test validity than misfitting outfit statistics, item fit was defined as an infit MnSq <1.4 in combination with a z-value of <2(124). A total number of >95% item fit was considered as evidence to support fit to the model and unidimensionality.

Principal component analysis of standardized residuals is usually the last step when investigating unidimensionality(125). The variance explained by measure; i.e., the dimension of the latent trait, is analyzed together with the variance in the first contrast. A high variance in the first contrast indicates a possible second dimension, and the principal component plot should be explored for a possible identification of the nature of this second dimension. In Study IV, 3%–5% of variance in the first contrast was considered very good, 5%–10% was considered good, and 10%–15% was considered fair(126).

3.6.5.4 Targeting

Targeting was examined through visual inspection of the person–item map, to make sure that the item difficulty covered all ability levels of the persons, and that floor and ceiling effects were absent(126). A maximum difference of 0.5 logits between person measure mean and item measure mean (by definition 0) was also considered as good targeting(126).

3.6.5.5 Internal consistency

In Study IV, item and person reliability, and person strata were examined. Item reliability refers to the ability to order the test items along the latent variable in a consistent way(127). The higher the item reliability coefficient is, the more confident we can be that the same hierarchy would appear across another sample. Low item reliability indicates that your sample size is not big enough to precisely locate the items along the latent trait. Person reliability refers to the ability to order the persons along the latent variable in a consistent way. A low person reliability implies that the test instrument may not be sensitive enough to separate between high and low performers, and more test items may be needed. Person reliability is analogous to the classical 'test reliability' and can be interpreted in the same way as Cronbach´s alpha.

Analogous to the Cronbach´s alpha, item and person reliability coefficients are in the range 0–1. When these coefficient values approach 1, they can be hard to interpret, and an alternative way is to report separation index or strata. Based on the reliability coefficient, separation index and strata can be calculated. Separation index refers to how many statistically different measurement levels there is in your sample, when very high and very low measures are removed from the calculation. A person separation index of 1.5 represents an acceptable level of separation, and a value >2 is generally considered a minimum requirement for satisfactory discrimination. Person strata is an alternative way to describe distinct person ability levels. Here, high and low measures are included. Person strata can be calculated using the formula = (4 * Separation index + 1) / 3(128).

Reliability coefficient values greater than 0.80 were considered good, and values greater than 0.90 were considered excellent.(126) Person strata were considered good when >3, very good when between 4–5, and excellent when >5(126).

3.7 ETHICAL CONSIDERATIONS

All studies in this thesis were approved by the appropriate Ethical Review board (Studies I–III the Regional Ethical review board, and Study IV the Ethical Review agency). Study I; 2005/849-31/3 and 2014-1359-32, Study II; 2017/983-31/1, Study III; 2016/682-31/2, Study IV; 2016/682-31/2 and 2020-05271. The studies were all conducted in accordance with the Declaration of Helsinki. Participating in the studies was voluntary and participants were informed that they could withdraw at any time without consequences.

Consent to participate in Studies I–III were obtained by the participants and/or their parents. For Studies I and III, written as well as oral consent was obtained after the participants had been given written and oral information about the studies. For Study II, oral information/consent was considered sufficient by the research team, as well as by the Regional Ethical Review board, since data were collected from medical records only and did not involve extra effort from the participant. For Study IV, oral consent was obtained (after oral information) for those individuals where clinical DIS videos were scored for this study. For participants where DIS scores were reused from other studies (that had separate approvals from the Ethical review boards), no consent/information was given, in accordance with the decision from the Ethical Review agency.

When someone has decided to participate in a study, he or she obviously wants to contribute to science. There are some obvious disadvantages for participants in a study. First, it takes time. Children with chronic disorders already have a lot of scheduled hospital appointments, which need to be considered by the researcher. Second, participation may be uncomfortable, or even painful. As a researcher, it is important to define the project in a way that minimizes uncomfortable (or even dangerous) parts. When research involves children, special consideration, for example regarding the research environment, is often needed. The balance between disadvantages for the participant and the importance of the results of the study must be considered. If possible, I think it is important to give the participants individualized information about their disorder/condition during the research process, since that may outweigh possible disadvantages of participating in the study. In my experience this type of information, being able to speak for a prolonged period of time with an expert in the field was often asked for and appreciated by children/adolescents as well as their caregivers. When designing the studies in this thesis, these ethical aspects were considered.

BPTI is a transient condition with recurrent episodes of torticollis in infants and young children. Delayed motor development is not uncommon among individuals still suffering from BPTI. Parents are often worried about the etiology and long-term outcome. Many parents explained that the reason for their participation was to receive more information about their child, and, to contribute to the general knowledge of BPTI. At the first follow-up, the children were examined by a pediatric neurologist and a physiotherapist. The procedure is common in standard care, and normally not appreciated as uncomfortable. All families received direct feedback of the neurologic examinations and motor assessments. At the second follow-up, the children as well as their parents, had the possibility of receiving information about BPTI and other related conditions, such as migraine, from a pediatric neurologist. The second follow-up included analysis of candidate genes. Since children and adolescents normally do not appreciate blood samples, they could choose between a blood sample or saliva sample. When the study was

finished, all families received a written summary of the findings, and for those families where variants in candidate genes were detected, oral information was given as well.

Dystonia is a condition that often severely affects quality of life, knowledge regarding effective treatment is limited, and reliable and valid assessment scales for children and young adults are scarce. Studies II–IV involved participants with dystonia. Many participants explained that living with a rare disorder, about which information is lacking, made them want to contribute to science in general, and particularly to the knowledge of dystonia. During the research process, participants had the possibility of receiving personal information and advice from a pediatric physiotherapist and from a pediatric neurologist. In my opinion, this was very appreciated by the children and their families and made the travel to the hospital and the time spent in the motion analysis laboratory at Astrid Lindgren Children's Hospital worthwhile.

Data for Studies II and IV were collected either from medical records, existing videos, or from scorings according to the assessment scale DIS in earlier studies. Here, ethical aspects mainly regard secure research storage. Another important ethical aspect of Study II was to present research results without revealing the identity of the participants. Attached to the published article from Study II, there was a video pre- and post-DBS of a young girl. Specific consent was signed by the girl as well as her caregiver. That girl had a very good effect from the surgical treatment and wanted other children with the same condition to receive that information.

For Study III, participating in the study included one interview, one examination, and two sessions with video recording. All these procedures are normal parts of standard care, and when performed in a respectful way, not uncomfortable. During video recording, the child/adolescent/young adult was asked to wear only shorts (and females a top as well). This may of course be inconvenient, but great efforts were made during the process to make participants comfortable. After examination/filming, the families received individualized advice from the pediatric physiotherapist and the pediatric neurologist.

4 RESULTS AND DISCUSSION

4.1 STUDY I: GENOTYPE, PHENOTYPE, AND LONG-TERM DEVELOPMENT IN BPTI

Study I was a prospective longitudinal follow-up case series of children with BPTI, with two onsite follow-ups. Thirteen children with BPTI met the inclusion criteria of the study, but one boy did not want to participate, and another boy was lost to the second follow-up. The first follow-up included eight girls and four boys (mean age 4 years, SD 2 years and 5 months). The mean age in the second follow-up was 13 years and 9 months, SD 2 years and 2 months.

The episodes of torticollis started at a median age of 2 months, and occurred at a median frequency once monthly, and ended at a median age of 21 months, Table 4. The duration of episodes was varying, and in 11 out of 12 children it was described as somewhere between 1 to 14 days. Our cohort resembles those of other children with BPTI previously reported regarding age at onset, duration, and frequency of episodes, as well as the time when the episodes resolve(32-34, 37). Eight of our 12 children were girls, and female dominance has been reported previously(34).

Table 4. Clinical features of our 12 children with benign paroxysmal torticollis

Patient number	Gender (F/M)	Age at onset (mo)	Attacks per month	Duration of attack(d)	Age at recovery	Associated symptoms	Involvement of alternating sides	Special tests, all normal
1	F	2	0.5	4–7	18mo	pallor, passive, vomiting	yes	EEG, CT, spinal Xray, LP
2	M	8	1	<7	3y	pallor, passive, vomiting, ataxia	yes	EEG, CT, spinal Xray
3	F	1	1	7–10	12mo	none	no	CT
4	F	12	1	3–7	2y	vomiting, irritability motor function impairment	yes	MRI, CT
5	F	1	0.3	7–21(28)	2y	pallor, passive, vomiting, photophobia	yes	CT
6	F	2	1	7	8mo	none	no	-
7	M	19	0.5	1	4y	vomiting, ataxia, vertigo	yes	EEG, CT
8	F	2	1	10–14	18mo	none	no	EEG
9	F	2	0.8	14	3y	vomiting	yes	-
10	M	1	1	5–10	10mo	none	yes	-
11	M	2	2	2–10	12mo	vomiting	yes	-
12	F	5	2	5	2 y	none	yes	-

M =male, F =female, mo =months, d =days, y =years, EEG =electroencephalogram, LP =lumbar puncture, CT =computer tomography of the brain, (–) indicates assessments not being performed. Slightly modified from ref: DevMed 2018; 60:1251-1255, doi: 10.1111/dmcn.13939

All motor assessments were normal in the first follow-up, and all children attended regular school with no report of any motor problems at the second follow-up. Five children out of 11 had developed either migraine, cyclic vomiting, or abdominal migraine at the second follow-up. All of them were reported to have rather mild symptoms, with no need of migraine-specific drugs or prophylactic treatment. No one had developed paroxysmal tonic upgaze, benign paroxysmal vertigo, paroxysmal dyskinesia, epilepsy, alternating hemiplegia, or episodic ataxia. Five children had at least one first-degree relative with migraine, but there was no family history of other paroxysmal disorders. Furthermore, no known mutations in candidate genes were found.

Altogether, these results support that BPTI is a benign condition that does not lead to neurological sequelae. Some findings though, need to be discussed more in detail and compared with previous reports and our current knowledge of the condition.

4.1.1 Nature of episodes

Although BPTI was described in 1969, diagnostic criteria were first formulated in 2013 by the International Headache Society, in the International Classification of Headache Disorders 3rd edition beta, ICHD-3 beta(32, 129). According to the ICHD-3 from 2018, BPTI is a condition with recurrent attacks of head tilt, with at least one of the following associated symptoms; pallor, irritability, malaise, vomiting, or ataxia(130). In our case series though, five out of twelve children did not have any associated symptoms. Despite this, we consider the BPTI diagnosis to be correct, since all suffered from recurrent episodes of torticollis, with normal motor function between attacks, and unremarkable clinical workups. This lack of associated symptoms has likewise been reported previously, indicating that such absence does not exclude a BPTI diagnosis(32-34).

No ocular symptoms were described in our cohort, and that is somewhat surprising, since they are often described as preceding a torticollis episode(14). One might speculate that associated symptoms and ocular symptoms, might be predictors of a less benign course of BPTI.

4.1.2 Family history

Together with benign paroxysmal vertigo (BPV), cyclic vomiting and abdominal migraine, BPTI is a disorder referred to as an "episodic syndrome that may be associated with migraine"(130). Individuals with these disorders are known to have a high prevalence of migraine in their family history, and they are likely to either have, or to develop, migraine(39). Furthermore, some individuals suffering from BPTI are later reported to develop BPV, cyclic vomiting, or abdominal migraine(32, 37-39). These disorders, as well as other 'less benign' paroxysmal disorders, have been described in the extended family, and there is evidence of a common genetic etiology in some cases(23, 42). Episodic ataxia, epilepsy, paroxysmal tonic upgaze, paroxysmal dyskinesia, alternating hemiplegia of childhood, and hemiplegic migraine are examples of such disorders.

With this said, we expected to find a high prevalence of migraine, and perhaps other paroxysmal disorders as well, in the family history of our children with BPTI. Indeed, five

of the 11 children had at least one first-degree relative who suffered from migraine, whereof one had hemiplegic migraine. Furthermore, one girl had a younger brother with BPTI. But apart from that, no other paroxysmal disorder, such as the aforementioned episodic ataxia, epilepsy, paroxysmal tonic upgaze, paroxysmal dyskinesia, alternating hemiplegia, and hemiplegic migraine, was reported. One might speculate about what the absence of these 'less benign' paroxysmal disorders in the family history means to our cohort of children. It may be possible that the outcome after BPTI is variable, and less benign in individuals with a family history of paroxysmal disorders other than non-hemiplegic migraine.

4.1.3 Development of paroxysmal disorders

We expected the prevalence of migraine to be higher in our cohort of children, compared to the general population. The prevalence of migraine is age-dependent and in adolescence prevalence ranges between 3%–19%(131, 132). At the second follow-up, with a mean age of 13 years and 9 months, 3/11 children suffered from migraine and one from abdominal migraine. Additionally, the child with abdominal migraine described transient symptoms consistent with a diagnosis of cyclic vomiting. Finally, one child had recovered from abdominal migraine. Children with migraine at the second follow-up reported mild symptoms, with no need of migraine-specific drugs during attacks or prophylactic treatment. These findings indicate that although migraines seem to appear more commonly after BPTI than in the general population, this may be a mild variant.

4.1.4 Genetic findings

Mutations in *CACNA1A* and *PRRT2* have been described in BPTI(38, 40-43). There are reports about families where several members harbor the same mutation; i.e. share a specific genotype, but display different phenotypes, such as BPTI, BPV, episodic ataxia, paroxysmal tonic upgaze, paroxysmal dyskinesia, hemiplegic migraine, or epilepsy(23, 42, 44, 133). Hemiplegic migraine is known to be caused by autosomal dominant mutations in *CACNA1A*, *ATP1A2*, *SCN1A*, and *PRRT2*. All these genes encode proteins that are involved in cell signaling, and mutations in these genes are known to cause other paroxysmal disorders as well. Consequently, we decided to analyze these four candidate genes in our BPTI population.

However, we did not find any known mutations in tested candidate genes. In two girls, however, variants of unclear clinical significance were found. In patient no1, a variant was detected in the *CACNA1A* gene (NM_023035) c.5176G>A (p.Val1726Met). In patient no12, a variant in the *ATP1A2* gene (NM_000702), c.2273G>C (p.Gly758Ala) was found, which has earlier been described in hemiplegic migraine. This patient had a brother with BPTI and a father with migraine, but they declined genetic testing. We believe that this variant may have been disease-causing. Neither of these girls had migraine or other paroxysmal disorders at the second follow-up.

Although mutations in *CACNA1A* and *PRRT2* earlier have been described in BPTI, the prevalence of such mutations in BPTI is not known(134). Shin et al. analyzed *CACNA1A* in eight children with BPTI to address that question. Similar to our study, they did not find any

mutations in their cohort. They speculated that mutations in *CACNA1A* are more likely to be found when there is a positive family history for other paroxysmal disorders than migraine, and our findings support that hypothesis. In addition, a study from 2018 reported that the presence of two or more different paroxysmal disorders in the family history was significantly associated with a *CACNA1A* mutation, in individuals with BPTI, BPV, or benign tonic upgaze(135). It is likely, that the absence of a positive family history of 'less benign' paroxysmal disorders in our cohort explains the lack of known mutations in candidate genes.

4.1.5 Long-term development: motor function and cognitive abilities

Gross and fine motor delay have been reported in BPTI, but since long-term follow-up studies are lacking, it is not known whether this is persistent (34, 36). In a case report from 2009, 5 out of 10 children with BPTI had gross motor delay during the period with episodes of torticollis(34). Similar results were reported in 2015, where two out of three children with BPTI had mild gross motor delay(36).

Cognitive dysfunction has been reported in BPTI, but it is not known whether this is more common than in the general population(135). In 2018, 22 individuals with BPTI were assessed regarding cognitive function, some of them formally tested and some of them evaluated via parents and medical records(135). No child fulfilled a diagnosis of intellectual disability, although three were considered to have cognitive dysfunction. The definition of cognitive dysfunction in this study was quite broad, including, for example, hyperactivity, attention deficiency, and the need for extra support in school, and accordingly comparison with the general population is not straightforward. Interestingly, two out of the three children with cognitive dysfunction had mutations in *CACNA1A*(135). Such mutations are known to be associated with cognitive impairment(136, 137). At the same time, two other children in this study, without cognitive dysfunction, had mutations in *CACNA1A*(135).

In our present study, we did not formally assess motor development during the period of recurrent torticollis, but three out of 12 were considered delayed by their parents and/or a physiotherapist. At the first follow-up, however, all motor assessments were normal. Furthermore, neurological examinations were normal in all children, except one child, who was considered generally hypotonic. At the second follow-up, with a mean age of 13 years and nine months (SD 2 years 2 months), no child had motor problems according to their parents. Accordingly, our data indicate that motor delay, if present, does not persist. In addition, all children attended regular school and no cognitive dysfunction was reported.

Our study contributes to the understanding of BPTI, and the findings indicate that BPTI is a benign condition that does not lead to neurological sequelae. Although migraine seem to develop more commonly than in the general population, there is a good chance that it is a mild variant. However, our cohort did not include paroxysmal disorders in the family history other than migraine, hemiplegic in one, and there was a lack of known mutations in candidate genes. Furthermore, during the torticollis episodes, several children did not display associated symptoms, and no one had ocular symptoms. In those cases, specifically if the child later develops a more severe migraine or other paroxysmal disorders, we encourage following the child and to consider genetic testing since the outcome may be less benign.

4.1.6 Strengths and limitations

A strength of this study is its longitudinal follow-up of more than a decade, with only one individual lost to the second follow-up. Moreover, different methods were used to answer research questions, such as physical examinations by a pediatric neurologist, motor assessments by a pediatric physiotherapist, evaluation of medical records and structured interviews of the children and/or parents by a pediatric neurologist, and finally, genetic testing of the children. There are some limitations of this study. First, no formal cognitive testing was performed, and second, family members could not be assessed for the sequence variant in *ATP1A2*. Another limitation is the relatively small number of children enrolled, but that is due to the rareness of the disorder, and most other reports include even fewer individuals.

4.2 STUDY II: RESPONSE TO DEEP BRAIN STIMULATION IN DYT-THAP1 DYSTONIA

This retrospective follow-up study was a case series of 14 individuals with DYT-THAP1-dystonia, from 5 European centers (Stockholm, Sweden; Milan, Italy; Montpellier, France; London, England; and Cologne, Germany), who had received pallidal deep brain stimulation (GPi-DBS). At the time of surgery, the median age was 18 years (range 8–57 years), and the median disease duration was 9 years (range 2–19 years). The median follow-up time after surgery was 4 years and 10 months (range 7 months–16 years). Before surgery, all individuals had generalized or segmental dystonia. Orolaryngeal dystonia causing speech impairment was present in all individuals, and swallowing was affected in 6. Three had documented musculoskeletal contractures, two in the neck and one in the feet.

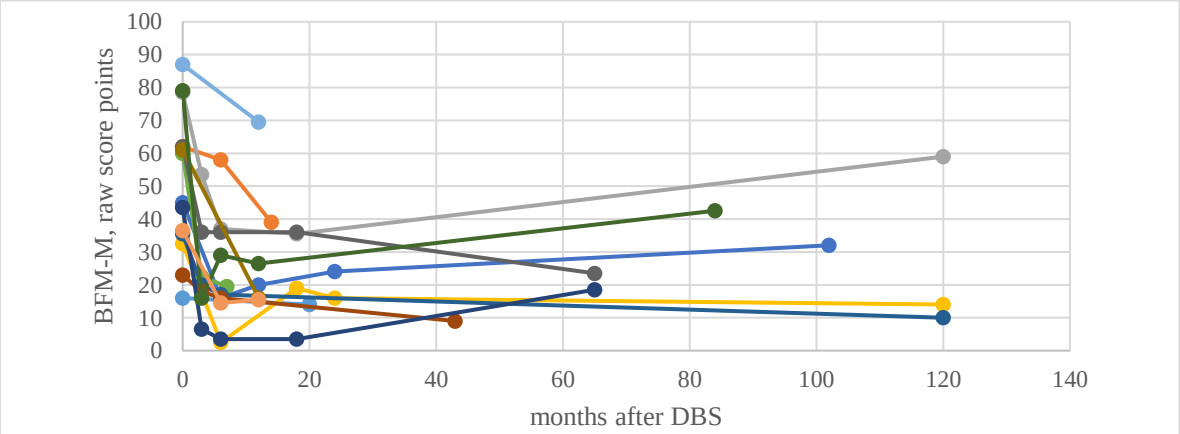
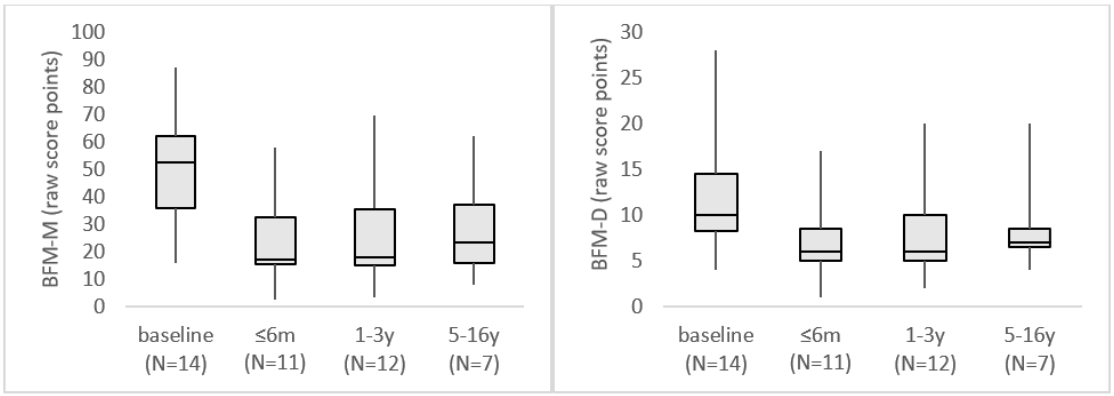


Figure 3. Spaghetti plot displaying 14 individuals with DYT-THAP1 dystonia, followed longitudinally after deep brain stimulation and evaluated by BFM-M. Please note that only 10 years of follow-up are included. Ref: J. Clin. Med. 2019, 8, 2163; doi:10.3390/jcm8122163

A positive effect was seen in all participants (Figure 3). The maximal effect was observed after a median time of 10 months, and the improvement remained stable for most individuals. The follow-up program differed between the five centers, but all participants were assessed pre- and postoperatively with BFM-M and BFM-D (Figure 4).

The BFM-M score was reduced from a median of 52.5 (IQR 35.8–62.0) before surgery to 18.5 (IQR14.4–37.3) at the last follow-up (p =0.001). The BFM-D score was reduced from a median of 10.0 (IQR 8.3–14.5) before DBS to 7.0 (IQR 5.3–9.5) at the last follow-up (p =0.006). Effect size at last follow-up was large, as assessed by the BFM-M (0.62) as well as the BFM-D (0.52).



(a) **(b)**

Figure 4. Boxplots showing BFM-M/D raw score points, with whiskers from minimum to maximum score. N = the number of individuals for whom data were available at each time-interval. (**a**) BFM-M boxplots at baseline and during follow-up at three different time intervals. (**b**) BFM-D boxplots at baseline and during follow-up at three different time intervals. BFM-M =Burke–Fahn–Marsden movement subscale, BFM-D =Burke–Fahn–Marsden disability subscale. Ref: J. Clin. Med. 2019, 8, 2163; doi:10.3390/jcm8122163

Side effects and complications were few. No intracranial bleedings or infections were reported. A wound infection was reported in one individual and one individual was successfully treated with corticosteroids due to edema around the lead tracks. Two individuals needed intracerebral revisions, in one case due to progressive worsening of symptoms, where two new electrodes were added, and in the other case due to rapid loss of function where one electrode was replaced.

DBS was first introduced in 1987 and was initially used to treat tremor and Parkinson's disease, and later increasingly used in other conditions such as dystonia and obsessive-compulsive disorders. During the past decade, outcome of pallidal DBS has been reported in different variants of dystonia and our knowledge has increased regarding the likelihood of an individual being a DBS-responder or not. The variable response to DBS is hypothesized to have multiple reasons, not only the etiology and the phenotype of the dystonia, but also the timing of the surgery, the anatomical placement of electrodes, and the stimulation parameters.

Although there are more than 20 years of experience with DBS in dystonia, it is still a challenge to understand the variable responses to the treatment in different individuals and there is a need to identify predictors for a good outcome after surgery. In the present study, we aimed to explore whether DYT-THAP1 dystonia is a predictor for good surgical outcomes and to what extent the timing of surgery correlates to the outcome. Furthermore, we have examined to what extent orolaryngel dystonia in DYT-THAP1 dystonia responds to DBS.

Based on the findings in this present study, some aspects explaining the variable effect in dystonia reduction after DBS need to be discussed more in detail. Together with previous knowledge regarding the effect of DBS in dystonia, these findings can guide clinicians in selecting individuals appropriate for surgery. Furthermore, some thoughts regarding possible new targets for DBS in DYT-THAP1 dystonia will be presented.

4.2.1 Aspects of genotype

In this study, we could demonstrate a good effect after GPi-DBS in individuals with dystonia caused by mutations in the *THAP1*-gene. Furthermore, the effect remained stable during a median follow-up time of about five years. It has already been described that at least some of the variable response to DBS in individuals with dystonia can be explained by genetic etiology(138-140). Extensive data exist on the outcome of DBS in DYT-TOR1A dystonia, which is often described as remarkably good and long-lasting(52, 54, 141). As an example, a long-term follow-up study including 47 individuals with DYT-TOR1A dystonia, showed that dystonia severity was reduced by an average of about 80% two years after DBS and in some patients, the effect remained stable during follow-up for up to 8 years(142). Other monogenic dystonia with reported good outcomes after DBS include DYT-KMT2B, DYT-SGCE, Lesh-Nyhan syndrome, and dystonia caused by *GNAO1*-gene mutations, whereas individuals with DYT-ATP1A3 dystonia or individuals with Glutaric aciduria type 1 do not seem to respond to the treatment at all(143-146).

Before the present study, only case reports or small case series mostly including 2–4 individuals with DYT-THAP1 dystonia and DBS treatment had been reported, and the scientific evidence to support DBS was scarce(49, 55-57). In 2015, however, a study reported results from 9 individuals with DYT-TOR1A dystonia and 8 with DYT-THAP1 dystonia, with a post-operative follow-up between 22 and 92 months, showing similar long-term improvement of BFM-M in DYT-TOR1A and DYT-THAP1 mutation carriers (−44% and −42% respectively)(54). Despite this, the outcome after DBS in DYT-THAP1 dystonia has been considered less beneficial compared to the case of DYT-TOR1A dystonia and some researchers have suggested genetic testing before decision making, and the finding of a DYT-THAP1 mutation could serve as an argument against surgery due to its lower probability of success(52, 54). Our present study, however, confirms that DBS is a safe and efficient treatment option for individuals with DYT-THAP1 dystonia. We agree that comprehensive genetic testing should be performed preoperatively. According to our findings, carriers of *THAP1*-mutations would then be considered well-suited for surgery.

4.2.2 Aspects of phenotype

In this study, all individuals had DYT-THAP1 dystonia, which is a disorder where dystonia is the only symptom; namely isolated dystonia. Individuals with isolated dystonia typically have a better outcome after DBS compared to individuals with dystonia combined with other neurological symptoms or individuals with lesions visible on an MRI scan of the brain(29, 80, 81). As an example, one review article with data from 76 children with different variants of dystonia, has shown that the overall improvement after DBS was 43.8 +/−31% (mean

+/−SD) when assessed by the BFM-M, but the likelihood of an improvement of more than 50% was higher in isolated dystonia (56%) compared to other dystonia variants (21%)(29). Similarly, a meta-analysis of DBS in pediatric dystonia showed a median improvement in BFM-M of 76% in inherited isolated dystonia (n =111), 26% improvement in inherited dystonia with degeneration or structural lesions (n =50), and 11% improvement in acquired dystonia (n =59)(81). Our study was in line with earlier findings, regarding good clinical outcomes in isolated dystonia after GPi-DBS, with a 58% median reduction of dystonia when assessed with BFM-M at the last follow-up after a median time of about 5 years.

Individuals with DYT-THAP1 dystonia typically have prominent orolaryngeal and craniocervical dystonia. The anatomical distribution of the dystonia is an important phenotypic aspect that may influence the outcome after DBS. A poor effect of DBS on orolaryngeal dystonia, causing problems with speech and swallowing, has been shown in DYT-THAP-1 dystonia as well as in KMT2B dystonia(49, 55, 57, 143). Our study could confirm those findings, since only 4 out of 14 individuals had some improvement of speech and swallowing. Overall, the dystonia reduction was more obvious in the trunk and limbs, compared to the orolaryngeal and craniocervical regions.

4.2.3 Aspects of location of leads and stimulation parameters

In this study, all individuals were treated with pallidal deep brain stimulation (GPi-DBS). This anatomical target is by far the most commonly-used one when treating dystonia, even though there are some reports regarding stimulation of the subthalamic nucleus (STN) or of the ventralis intermediate nucleus (VIM) of the thalamus(29, 146-150). Due to the limited number of cases, comparison between effect is not straightforward. It is possible, however, that some individuals with dystonia would benefit more from other anatomical DBS targets than GPi. To address this, it is possible to compare different targets intraoperatively, or as recently suggested, during a whole week in an inpatient Neuromodulation Monitoring Unit, with electrodes at multiple possible targets in the basal ganglia and thalamus(151).

Recently, VIM-DBS was reported to be successful in the treatment of laryngeal dystonia, which is interesting due to contradictory reports regarding the effect of GPi-DBS on such dystonia(152). Another study has shown a similar effect in DYT-THAP1 dystonia, where a dramatic response on laryngeal dystonia was reported after VIM-DBS, where GPi-DBS had failed(153). There is some evidence that the target thalamus could be used to treat orolaryngeal dystonia regardless of etiology(152, 154). It would be interesting in future studies to examine the VIM target for DYT-THAP1 patients with severe orolaryngeal dystonia.

In our study, stimulation parameters were chosen according to the local traditions at each center. Stimulation parameters at last follow-up were reported for 13 out of 14 individuals, and were as follows; 8 patients had stimulation frequency of 130 Hz, 2 had 180 Hz and the three remaining had 100-125 Hz. The choice of stimulation parameters is known to affect DBS outcome, but that was not examined in our study(146).

4.2.4 Aspects of timing of surgery

In this study, there was no correlation between age at surgery, disease duration before surgery, or disease burden at the time of surgery and the effect of DBS. In our study, individuals with preoperative contractures were all responders to the DBS, with 68%, 78%, and 46% reductions of dystonia at last follow-up, respectively, as assessed by the BFM-M score. The timing of DBS has been evaluated before in several reports and review articles and positive predictive factors, such as a shorter preoperative duration of dystonia, a shorter proportion of life with disease, a younger age at surgery and no fixed preoperative skeleton contractures, has been proposed(77, 146, 155, 156).

Our limited number of patients does not allow for definitive conclusions regarding possible outcome predictors for timing of DBS surgery, but nevertheless, the data support offering pallidal DBS to all individuals with pharmaco-resistant DYT-THAP1 dystonia, regardless of age, disease duration, or the presence of fixed deformities. However, due to the rapid and long-lasting effect of DBS treatment, intervention should be considered early in children and young adults, to prevent unnecessary morbidity. Another reason to consider DBS early in the disease course, is that DBS might have a long-term neuroprotective effect. In studies of individuals with Parkinson's disease, it has been hypothesized that DBS might slow the degeneration of dopaminergic neurons in the Substantia nigra, and furthermore, in some Parkinsonian rat models, STN-DBS has been shown to increase the survival of dopamine neurons in the Substantia nigra pars compacta(79).

4.2.5 Strengths and limitations

A strength of this study was the international collaboration, with centers from five European countries (Stockholm, London, Montpellier, Cologne, and Milan) participating. Together the co-authors of this study had a broad clinical experience and deep preclinical knowledge, since there were neurosurgeons, pediatric neurologists, neurologists, and geneticists contributing to the study. There are several limitations of this study, however. Due to the rareness of the DYT-THAP1 dystonia disorder, there was a limited number of included individuals. Another obvious limitation of this study is that it is a case series, without a control group, and as such it yields a low level of evidence regarding treatment effect compared to a randomized controlled study.

Risk of confounding bias: It is possible that the participants improved for other reasons than the GPi-DBS. First, a spontaneous improvement (regression towards the mean) is possible, but is not likely, due to our knowledge about the natural course of DYT-THAP1 dystonia. Second, changes in the pharmacological treatment could theoretically lead to improvement. We did not systematically check changes in the treatment during the follow-up, but all participants had received multiple pharmacological treatments before the DBS, without a significant or long-lasting effect. Third, improvements may be caused by a placebo effect. A placebo effect is common after a surgical procedure, but the effect often diminishes over time. Our follow-up was very long (median time of 5 years, range 7 months to 16 years), supporting our conclusions of the effects after DBS. The fact that all individuals were evaluated with a dystonia rating scale would also reduce the risk of a placebo effect, even though the rater was not blinded.

Risk of selection bias: The five contributing centers included all known individuals with DYT-THAP1 dystonia with a follow-up of more than 6 months. Nonetheless, there is still a risk of selection bias that may affect the possibility of generalizing our findings to the entire population of individuals with DYT-THAP1 dystonia. It is possible that some individuals were considered unsuitable for treatment in the first place, and therefore not offered the DBS surgery. Individuals with a severe clinical picture with a long proportion of life with dystonia, or individuals with a poor socioeconomic status might have been considered less suitable. Another aspect that may lead to selection bias in our study, is that the pathogenicity of the THAP1-variants was not systematically confirmed through functional analysis. There is a possibility that some of the participants did have a rather benign sequence variant in the *THAP-1* gene. Our research group, with senior geneticists and neurologists participating, discussed each participant thoroughly. Specifically, patients no 1, 4, 6 and 8 were discussed, due to less-available genetic information. When clinical data regarding phenotype, family history, and genetic findings were summarized, a diagnosis of DYT-THAP1 dystonia for each participant was considered appropriate by the research group.

Risk of information bias: In this study there is a risk for some information bias. First, there is the risk of misclassification. Each participant was rated with BFM by a clinician, which may lead to an observer bias (were the clinicians familiar enough with BFM? did they have preconceived notions?). Since the clinicians who rated the participants were working in a center where BFM is a common evaluation tool, they probably were familiar with the BFM. A serious disadvantage however, was that the rater was not blinded. Second, a lack of accurate measurement keys may lead to information bias. In this study, the only outcome measure was BFM-M and BFM-D. No evaluations regarding quality of life or participation were performed, and dystonia reduction does not necessarily lead to a higher quality of life or a higher degree of participation(157). Third, there is the risk of recall bias. In any retrospective study, there is a risk of recall bias. That risk is quite low in our study however, since the information was gathered from medical records. Furthermore, we had a well-defined protocol for data collection from medical records. All participating center had their own follow-up procedures, however, and during data analysis this needed to be considered.

4.3 STUDY III: RELIABILITY AND VALIDITY OF THE DIS IN INHERITED DYSTONIA

Study III was an instrument evaluation, a cross-sectional, two-center study of the DIS in inherited or idiopathic dystonia. Eleven males and 9 females with inherited or idiopathic dystonia from Stockholm, Sweden (n =14) and Leuven, Belgium (n =6), with a median age of 16 years and 7 months, range 6 to 24 years, were included in the study. All completed a first visit, and 19 were assessed at a second visit within a time frame ranging from a minimum of 2 hours to maximum of 2 days between visits. One girl was not invited to the second visit, since the investigators considered her to be uncomfortable during the first visit.

Inter-rater reliability: The DIS, the DIS-D, and the DIS-CA showed moderate-to-good inter-rater reliability, with ICC values of 0.83, 0.87, and 0.71 respectively. For the body regions, the inter-rater reliability was lower, ranging from poor to moderate/good, with ICC values ranging between 0.10–0.80, with most values between 0.6 and 0.7.

Test–retest reliability: The DIS, the DIS-D, and the DIS-CA showed good-to-excellent test–retest reliability with ICC values of 0.95, 0.88, and 0.93 respectively. For the body regions, ICC-values were lower, comparable with the inter-rater reliability, ranging from poor to good. The SEM and MDD for the DIS were 3.98% and 11.04%. For the DIS-D, 8.16% and 22.61% and for the DIS-CA, 3.80% and 10.54%, respectively.

Concurrent validity: The Spearman correlation coefficient between the DIS-D and BFM-M was 0.88 ($p < 0.01$).

Altogether, these results indicates that the DIS might be a valuable assessment tool for evaluation of dystonia and choreoathetosis in inherited and idiopathic dystonia conditions, but some findings need to be discussed more thoroughly.

4.3.1 Inter-rater reliability

In our population with inherited/idiopathic dystonia, inter-rater reliability was moderate-to-good for the DIS, the DIS-D, and DIS-CA. To interpret this more thoroughly and to put it in context, we can compare these results with previous publications regarding the DIS in the dyskinetic CP population. Another possibility is to compare the results from the DIS-D with reliability studies on other dystonia rating scales, such as the BFM-M, UDRS, Movement Disorder Childhood-Rating Scale or the BADS.

4.3.1.1 Comparison with inter-rater reliability for the DIS in dyskinetic CP

In 2012, in the original publication of the DIS (n =25), inter-rater reliability for two independent raters was shown to be excellent for the DIS and the two subscales(88). Regarding body regions, moderate-to-good inter-rater reliability was found for most regions, although some were lower/higher. In the dystonia subscale, ICC-values for body regions and total values were generally lower during action than in rest. In the choreoathetosis subscale, there was no difference in ICC-values for body regions in action nor in rest.

In 2013, inter-rater reliability among raters who were unexperienced in the assessment of dystonia and choreoathetosis, but specifically trained in the DIS, was examined(158). Inter-rater reliability was examined for two junior physiotherapists and was found to be good for the DIS and the two subscales (n =25). For three senior physiotherapists, inter-rater reliability was moderate-to-excellent for the DIS and the two subscales. Regarding body regions, all physiotherapists had inter-rater reliability ICC values mostly ranging from poor-to-moderate, with some higher values between good and excellent.

DIS: In Study III, inter-rater reliability for the DIS was good, whereas it was good-to-excellent in studies regarding dyskinetic CP(88, 158). This indicates that the DIS may be as reliable for assessing dystonia/choreoathetosis in inherited/idiopathic dystonia as in dyskinetic CP.

DIS-D: In Study III, inter-rater reliability for the DIS-D was good, whereas it was moderate-to-excellent in studies regarding dyskinetic CP(88, 158). Body regions assessed during rest had lower inter-rater reliability in our present study compared to the original publication of the DIS, but higher inter-rater reliability compared to what was found in the study with inexperienced raters. Body regions assessed during action had similar inter-rater reliability compared to the original publication. This indicates that the DIS-D may be as reliable for evaluating dystonia in inherited/idiopathic dystonia as in dyskinetic CP.

DIS-CA: In Study III, inter-rater reliability for the DIS-CA was moderate, which was lower compared to previous publications in the case of dyskinetic CP(88, 158). In dyskinetic CP, the DIS-CA, as well as the body regions in the choreoathetosis subscale, had higher inter-rater reliability compared to the dystonia subscale, whereas our study showed the opposite. In the original article of the DIS, they speculated that the reason behind higher ICC-values for choreoathetosis, was that choreoathetosis is easier to recognize compared to dystonia due to "the lack of sustained postures in choreoathetosis and the more identifiable nature of choreoathetosis"(88).

In-depth analysis of raw data from our study was performed to understand the low inter-rater ICC values in the DIS-CA compared to the DIS-D. In a first step, absolute percentage agreement among the three raters was analyzed, but no obvious pattern could be found explaining low ICC-values. In a second step, variability of raw scores was analyzed, since low variability among raw scores is known to cause lower ICC-values(115). Here, differences were found between the two subscales. About half of the individuals in this study did not have choreoathetosis in action items, and accordingly scored "0" on many test items. Dystonia, on the other hand, was present in most individuals in action as well as in rest, leading to a higher variability in raw scores. Hence, this low variability in raw scores in the DIS-CA (and not disagreement among the raters) may at least partly explain the overall lower inter-rater reliability. This indicates that it is possible that the DIS-CA is as reliable in inherited and idiopathic dystonia as in dyskinetic CP, but future research is warranted with more individuals who have choreoathetosis.

Additional comments on the DIS-D during action: When analyzing variability of raw scores, it was obvious that for some actions (the eyes, neck, arm right distal, leg right distal) more than half of the subjects scored a "4" for dystonia; i.e., dystonia was present almost all of the time in a maximal range of motion in these body regions during action. This can probably explain lower ICC-values for these test items. Poor ICC-values were also seen in the trunk during action. Trunk was perceived difficult to score by the three raters, and disagreement among the raters was clearly an explanation for the low ICC values. In depth analysis of raw data showed that half of the study participants were unable to perform the trunk actions. Accordingly, the ICC values for trunk during action must be interpreted with caution.

4.3.1.2 Comparison with inter-rater reliability for other dystonia rating scales

The BFM-M was evaluated regarding inter-rater reliability by the developers in 1985, but only for the total score, and not the body regions(86). Here, ten patients with dystonia were evaluated by three raters, and the Spearman correlation coefficient was calculated, and ranged between 0.85 and 0.96.

The BADS was evaluated regarding inter-rater reliability by the developers in 1999, and here ten patients were evaluated by 13 raters(87). They reported that the ICC for the total score was 0.87, whereas for body regions the ICC ranged between 0.06 and 0.71, with a mean value of 0.42.

In 2003, Comella et al. evaluated inter-rater reliability for the BFM-M, the UDRS and the GDS(85). The reliability coefficient ICC was used for total scores, and the Kendall's coefficient of concordance and generalized weighted kappa was used for body regions(85). 103 patients with isolated dystonia were evaluated by 25 raters, but each rater only evaluated a subset of patients. In this study, similar results were obtained for all three scales. The ICC for the total score ranged between 0.71–0.78. Regarding body regions, the Kendall's coefficient of concordance ranged between 0.52–0.87, with most values between 0.6 and 0.7, and generalized weighted kappa ranged between 0.37–0.90, with most values between 0.7 and 0.8.

In 2010, Monbaliu et al. evaluated inter-rater reliability for the BFM-M, the BADS, and the UDRS in a population of children and young adults with dyskinetic CP(91). Ten individuals were assessed by three raters each. The ICC for total scores were 0.86, 0.87 and 0.79 respectively. ICC values for body regions were lower, ranging between 0.62–0.86, 0.39–0.76, and 0.37–0.78 respectively.

In conclusion, compared to these other dystonia assessment scales, the DIS has similar inter-rater reliability in the population of individuals with inherited and idiopathic dystonia, for total scores as well as for body regions. Overall, lower inter-rater reliability was seen for body regions compared to total scores in our study of DIS in inherited and idiopathic dystonia, as well as in the above-mentioned studies evaluating different dystonia rating scales.

4.3.2 Test–retest reliability

Good-to-excellent test–retest reliability was found for DIS, DIS-D and DIS-CA, with ICC-values of 0.95, 0.88 and 0.93 respectively. This is in line with previous studies in the dyskinetic CP population, confirming that the DIS produces stable results over time and can be used for repeated assessments(159). A high test–retest reliability indicates a high intra-rater reliability, although that was not analyzed in this study. The SEM and MDD values for the DIS were 3.98% and 11.04%, respectively. This means that a score difference of 11% is a true change, and not due to measurement error. For DIS-D, the SEM and MDD were 8.16% and 22.61%, respectively, which is similar to previous reports regarding the BFM-M scale (SEM 9.88% and MDD 27.39%)(91).

4.3.3 Concurrent validity

Today, most children with inherited/idiopathic dystonia participating in intervention studies are evaluated with the BFM-M/D dystonia rating scale(86). In the present study, concurrent validity for the DIS-D, when compared with the gold standard BFM-M, was 0.88 when assessed with the Spearmen correlation coefficient. For dyskinetic CP, similar results have been reported, concurrent validity for the dystonia subscale was 0.84 as assessed by the Pearson correlation coefficient and BADS used as a gold standard(87, 88). This indicates that DIS-D, as well as BFM-M, can be used in intervention studies for this population.

4.3.4 Advantages and disadvantages of the DIS

The DIS is a comprehensive assessment scale for dystonia and choreoathetosis designed for children and young adults with dyskinetic cerebral palsy. When the scale was developed, recommendations from the Taskforce on Childhood Movement Disorders regarding important components of assessment were considered. Accordingly, the DIS was designed to assess the duration/amplitude of dystonia/choreoathetosis, in action/rest separately, in multiple body regions (distinction between proximal and distal extremities) on a 5-point rating scale. This enables evaluation of dystonia and choreoathetosis according to their dominant features and in the context in which they occur.

One advantage of the DIS is that different outcome measures may be chosen, depending on your clinical situation or your research question. The total DIS outcome measure and the outcome measures for the two subscales, DIS-D and DIS-CA, are probably useful in all settings. With these outcomes, it is possible to calculate a ratio of dystonia and choreoathetosis in specific disorders, and thereby increase our knowledge regarding rare disorders. In intervention studies, it will be possible to examine the effect of these movement disorders separately. Furthermore, it may be possible to link DIS assessments with other findings, such as brain lesions in basal ganglia/cerebellum/cortex cerebri, thereby learning more about the pathophysiology behind these movement disorders. Several other outcome measures may be interesting, such as dystonia during action and rest, making it possible to describe the dystonia burden during rest and action separately.

Another advantage with the DIS is that each outcome measure of the DIS corresponds to impairment only. Dystonia and choreoathetosis represent impairments (corresponding to the ICF domain body function and structures), which interferes with activities and the child's ability to participate in different life situations(160). The aim with the DIS is to measure impairment, and not activity limitation or participation restriction. Most other dystonia rating scales, aiming to assess impairment, such as the BFM-M and BADS, have total scores, and scores from body regions, that are combinations of impairment and activity limitation. As an example from the BFM-M, when scoring the severity factor for leg dystonia, you consider whether assistance is needed for walking, or whether walking is impossible due to dystonia. The DIS assesses the presence and severity of dystonia and choreoathetosis in rest and in action, but if an individual has an activity limitation and cannot perform an action, that item is excluded from the assessment. When actions are excluded in the assessment, the maximum raw score of the DIS is reduced. Therefore, the result of the DIS is mostly presented as a percentage score, where only rest postures and actions performed by the individual are included. Accordingly,

only impairment is assessed in the DIS. In intervention studies, aiming to reduce the amount of dystonia/choreoathetosis, and perhaps also aiming to improve function or quality of life, other assessment scales must be used as a complement to the DIS.

Due to the above-mentioned advantages, our research group aimed to examine whether the DIS could produce reliable and valid outcome measures in a population with inherited or idiopathic dystonia. Our study showed good-to-excellent inter-rater reliability, test–retest reliability and concurrent validity for the DIS and the DIS-D, and moderate-to-excellent inter-rater reliability and test–retest reliability for the DIS-CA. Accordingly, we concluded that the DIS might be a promising tool for evaluating dystonia and choreoathetosis in inherited and idiopathic dystonia.

In the paragraphs above, advantages of the DIS have been highlighted. Furthermore, the possibility to use the scale in a population outside the one it was developed for has been discussed. However, the DIS has some disadvantages as well. First, the DIS is comprehensive, and therefore time-consuming to administer and to score. Due to this, the DIS is often disregarded in a clinical setting. Another obvious disadvantage of the DIS is that comparison of DIS scores (the outcome measure) is not straightforward. When comparing the sum of raw scores, or percentage scores, you need to consider the number of actions performed/not performed. The outcome measure of the DIS is relative to the number of actions being performed. To address these two problems, effort is currently being made to create a shorter, more clinically useful, version of the DIS, where the outcome measure is absolute, and not relative, thereby enabling comparison (Study IV).

4.3.5 Strengths and limitations

A strength of this study is the collaboration between KU Leuven and Karolinska Institutet, enabling inclusion of 20 individuals with rare disorders. Another strength is that three raters videotaped, examined, and interviewed all participants, and scored all videos in a time frame of four weeks. Furthermore, the raters had completed a DIS instructional course in Belgium and had broad clinical experience in dystonia and choreoathetosis. The study design allowed us to evaluate inter-rater reliability, test–retest reliability, and concurrent validity. Our findings indicate that the DIS might be a promising tool to evaluate dystonia and choreoathetosis in individuals with inherited or idiopathic dystonia.

However, a limitation of the study was that content or construct validity of the DIS in inherited or idiopathic dystonia was not examined. It is possible that the content and construct of the DIS is more suitable for some dystonia populations. As an example, actions included in the DIS are chosen to suit non-ambulatory individuals, since dyskinetic CP is known to cause severe motor problems. Standing and rolling are included in the DIS, whereas walking is not. It is possible that a higher proportion of individuals with isolated dystonia is ambulant, compared to individuals with dyskinetic CP or individuals with metabolic causes of dystonia. For individuals with a higher level of functioning, it might be advantageous to be able to assess the presence of dystonia during walking or other more complicated actions. It is possible, that although the DIS showed good psychometric properties in this study, individuals with isolated dystonia are better assessed using other scales.

Another limitation of this study is that the participants are heterogenous. It is possible, that if individuals with a more similar phenotype had been assessed separately, the inter-rater reliability and test–retest reliability would differ between the groups. It could have been possible to analyze individuals with isolated dystonia, and individuals with a mixed phenotype separately. One might speculate that individuals with isolated dystonia are easier to score, due to absence of choreoathetosis/spasticity. This would lead to higher ICC values in that group compared to individuals with a mixed phenotype. When generalizing the findings from this study to the population of children and young adults with inherited and idiopathic dystonia, this must be kept in mind. Due to the rareness of these disorders, however, we chose to analyze isolated and mixed phenotypes together.

Finally, another limitation of this study is its small sample size. Initially, we aimed at including 25 individuals, since the original DIS publication in dyskinetic CP had 25 participants. Due to the rarity of these disorders, inclusion of individuals was not easy, and based on power calculations (see method section) we chose to include only 20 individuals. However, individuals with severe dystonia are not able to perform all action items in the DIS, and when analyzing data, it was clear that trunk actions were performed by approximately only half of the participants. With more individuals included, more accurate results might have been found for those action items. Accordingly, further research with more participants might be warranted.

4.4 STUDY IV: THE DIS-II: DEVELOPMENT, CONSTRUCT VALIDITY, AND RELIABILITY

In Study IV, a cross-sectional, two-center study, the DIS-II was developed, and construct validity and reliability (internal consistency) were evaluated using Rasch measurement model analysis.

4.4.1 Development of the DIS-II

A preliminary DIS-II was developed by the research team during iterative discussions. Results from an online whole-day advisory expert meeting, clinical aspects, utility of the scale, scoring feasibility, and previous existing psychometric properties regarding the original DIS were considered during these discussions.

4.4.1.1 Advisory expert meeting

It was obvious that the attending experts were in favor of shortening the DIS. In the original DIS, 12 body regions are scored, and 69% of the attendees wanted to reduce that number. Furthermore, each body region is assessed in rest and in two actions, and 87% of the attendees wanted to delete one of those two actions. Finally, 93% wanted to reduce the number of scale steps in the rating scale, the original DIS being a 5-step scale.

4.4.1.2 Discussions in the research team

First, the number of body regions were reduced from 12 to 11. In the expert meeting, there was no obvious agreement for which body region should be removed. Based on the high number of individuals unable to perform trunk actions, and the previously reported low inter-rater reliability values, the research team decided to delete the trunk region(158, 161). Second, the number of actions were reduced from two to one per body region, Table 5. Here, the advisory expert group could guide the research group in the decisions, since there was a consensus regarding which action to remove for most body regions. As a result, the number of videos to record was reduced from 26 to 14 and the number of items to score was reduced from 144 to 88.

Table 5. Original Dyskinesia Impairment Scale versus Dyskinesia Impairment Scale-II

Body region	**Original DIS**		**DIS-II**	
	Action	*Rest*	*Action*	*Rest*
Eyes	Eye tracking Eye blinking	Eyes in sitting	Eye blinking	Eyes in sitting
Mouth	Mouth open/close Speech	Mouth in sitting	Speech	Mouth in sitting
Neck	Lateroflexion Rotation	Neck in sitting	Rotation	Neck in sitting
Trunk	Active sitting Forward flexion	Trunk in sitting	Trunk not included in DIS-II	
Arm proximal	Arm abduction Grasp/reach a pen (supine)	Arm proximal in sitting	Grasp/reach a pen (supine)	Arm proximal in sitting
Arm distal	Grasp/move a cup Grasp/move a pen	Arm distal in sitting	Grasp/move a cup	Arm distal in sitting
Leg proximal	Rolling Standing	Legs proximal in lying	Rolling	Legs proximal in lying
Leg distal	Rolling Heel/toe tapping	Legs distal in lying	Heel/toe tapping	Legs distal in lying
Rating scale for amplitude and duration	0 = no D/CA 1 = <10% D/CA 2 = ≥10-<50% D/CA 3 = ≥50-<90% D/CA 4 = ≥90% D/CA / = action not performed		0 = no D/CA 1 = <50% D/CA 2 = ≥50 D/CA 3 = action not performed	0 = no D/CA 1 = <50% D/CA 2 = ≥50 D/C

DIS =Dyskinesia Impairment Scale, D/CA =dystonia/choreoathetosis

The last step in the development of DIS-II was the rating scale. In the original DIS, the amplitude/duration of dystonia/choreoathetosis are assessed on a 5-step scoring scale, where 0 indicates no dystonia/choreoathetosis, and 4 indicates that dystonia/choreoathetosis is present >90% of the range of motion or of the time. Each body region is assessed during rest and action. Actions participants were note able to perform are deleted in the assessment (scored as a "/"), and the DIS result is expressed as a percentage score of items performed. To compare results, information regarding actions/rest postures being performed or not is needed.

When creating the DIS-II, the purpose was to make it shorter but also to make comparison easier. Consequently, the research team decided to include "not being able to perform/not done" in the rating scale for action items. From a clinical point of view, the research team identified two rating scale options. The first option had 4 steps for rest items and 5 steps for action items, whereas the second option had 3 steps for rest items and 4 steps for action items. Rasch measurement model analysis was used to select the most appropriate rating scale option.

4.4.2 Rasch analysis of the DIS-II

4.4.2.1 Rating scale functioning

First, the Rasch measurement model analysis was used to select the most appropriate rating scale option. The two rating scale options were analyzed separately. When comparing summary statistics, person reliability was 0.97 for both options, and item reliability 0.97 for the first option and 0.96 for the second option, i.e. almost identical for both options. When analyzing rating scale functioning though, obvious differences in favor for the second rating scale option were found, i.e, a rating scale with 3 steps for rest items and 4 steps for action items.

First, visual inspection of distinct peaks for each category was not possible with the first rating scale option, whereas distinct peaks were identified with the second rating scale option. Second, item step calibration was inconsistent in 16 out of 88 items with the first rating scale option (the thresholds were disordered), whereas in only 2 items with the second rating scale option. Therefore, the 3/4 step rating scale option was chosen for the DIS-II, and all further analyses are based on that option (Table 5).

4.4.2.2 Unidimensionality

Secondly, unidimensionality of the DIS-II was investigated with item point-measure correlations, fit statistics, and principal component analysis (PCA) of standardized residuals (Table 6).

Item point-measure correlations were all positive and ranged between 0.16–0.75, supporting unidimensionality.

Fit statistics for item showed >95% fit to the model in DIS-II and DIS-II-D/DIS-II-CA, hence evidence of goodness of fit to the model and unidimensionality. Only two test items were misfit in DIS-II, amplitude, and duration of choreoathetosis in the eye region during action.

Principal component analysis (PCA) of standardized residuals also supported unidimensionality (Table 6). PCA of the DIS-II showed a 45.1% variance explained by measures and 9.0% unexplained variance in the first contrast. For the DIS-II-D and DIS-II-CA, variance explained by measure was 49.6% and 41.1%, respectively, with 8.1% and 12.6% variance in the first contrast. An unexplained variance in the first contrast of 5%–10% was considered good, and 10%–15% was considered fair; i.e. good for DIS-II and DIS-II-D, and fair for DIS-II-CA.

Examination of the PCA factor plot explained the relatively high unexplained variance in the first contrast for DIS-II, DIS-II-D and DIS-II-CA. It was clear that a second dimension was action and rest, which was expected, since dystonia and choreoathetosis behaves differently during rest/action, and accordingly are assessed in both contexts. In DIS-II, a third dimension, dystonia and choreoathetosis was visible, which was not unexpected, since these movement disorders are defined differently. Therefore, the unexplained variance was not considered a problem.

4.4.2.3 Targeting

Targeting for DIS-II was investigated through visual inspection of the person–item bar chart, which showed a good spread of item difficulties covering all person abilities. Floor and ceiling effects were absent. Additionally, the difference between person measure mean and item measure mean in DIS-II, and DIS-II-D were –0,26 and 0.31 respectively, indicating good targeting. For DIS-II-CA, the person measure mean was –0,82 though, indicating that there might be an insufficient number of items covering persons with little choreoathetosis. This is explained by the fact that choreoathetosis was absent/almost absent in some individuals, and for those the DIS-II-CA is not suitable.

4.4.2.4 Internal consistency

Internal consistency, as assessed by item and person reliability and person strata, was excellent for the DIS-II and the two subscales DIS-II-D/CA, (Table 6). For DIS-II, person strata was 8.03, indicating that the scale can separate individuals into eight distinct ability levels.

Table 6. Rasch analysis results for the DIS-II, DIS-II-D and DIS-II-CA

	DIS-II	**DIS-II-D**	**DIS-II-CA**
Principal component analysis			
Variance explained by measures (%)	45.1	49.6	41.1
Unexplained variance in the first contrast (%)	9.0	8.1	12.6
Person			
Measure mean (SD)	-0.26 (1.06)	0.31 (1.46)	-0.82 (1.18)
Measure range	-4.13–2.41	-4.24–4.72	-6.10–1.70
Mean standard error	0.16	0.26	0.25
Reliability	0.97	0.96	0.93
Separation index	5.77	4.63	3.6
Strata	8.03	6.5	5.13
Item			
Measure mean (SD)	0.00 (0.73)	0.00 (0.56)	0.00 (0.53)
Measure range	-1.76–1.36	-1.31–1.24	-0.97–1.11
Mean standard error	0.13	0.15	0.14
Reliability	0.96	0.92	0.92
Separation index	5.06	3.42	3.46

DIS-II =Dyskinesia Impairment Scale-II, DIS-II-D =Dyskinesia Impairment Scale-II dystonia subscale, DIS-II-CA =Dyskinesia Impairment Scale-II choreoathetosis subscale, SD =standard deviation

4.4.3 Advantages and disadvantages of the DIS-II

The main objective of this study was to shorten the original DIS to make it more attractive for use in both research and clinic. In DIS-II, the number of videos to record was reduced from 26 to 14, and the number of items to score was reduced from 144 to 88. Furthermore, the rating scale has been simplified, since the number of scoring scale steps have been reduced from 5 in the original DIS, to 3 for rest items and 4 for action items. Altogether this reduces the maximum score from 576 to 220 points. Another objective was to make a comparison over time or between individuals easier. By including "not being able to perform/not done" in the rating scale for action items, the outcome measure of the DIS-II is absolute, and not relative as in the original DIS, thereby enabling comparison.

An important advantage with the DIS-II is that, in addition to a raw score, the outcome measure may be presented as a logit measure. Logit measures are superior to raw scores, since logit measures are interval data, whereas raw scores are ordinal data. With the Rasch measurement model analysis, conversion tables for raw scores to logit measures have been created for the DIS-II, DIS-II-D, and DIS-II-CA. It is possible to present other results from the DIS-II as well, for example dystonia during action or choreoathetosis during rest, but then raw scores must be used, since conversion tables were prioritized for the total DIS-II and the two subscales DIS-II-D and DIS-II-CA only.

Further advantages of the DIS-II include its good psychometric properties. The result from the Rasch measurement model analysis demonstrates evidence of construct validity of the DIS-II. Scores obtained from individuals with dyskinetic cerebral palsy, inherited, idiopathic, or acquired dystonia when evaluated with DIS-II according to a standardized video protocol, provide a valid measure of the amount of dystonia and/or choreoathetosis. The analysis of the rating scale functioning, the goodness-of-fit evaluation, together with the principal component analysis, show evidence of a unidimensional construct. Furthermore, the high person reliability indicates that the DIS-II is able to separate children and young adults with dystonia and choreoathetosis into eight distinct ability levels. Altogether, this implies that this new scale may be sensitive to change and thus is an appropriate tool to evaluate dystonia and choreoathetosis in clinical practice and in intervention studies.

A disadvantage of the DIS-II is that the trunk is not assessed at all. The trunk is assessed in the original DIS, as well as in the BFM, but was removed from the DIS-II. The reason behind this was that a high number of participants were not able to perform the trunk actions of the DIS, and furthermore, the research group as well as the attendees of the expert meeting considered trunk during rest complicated to score. The decision was not easy to make though, since dystonia and choreoathetosis in the trunk region may severely affect function, activity, participation, and quality of life.

Another disadvantage of the DIS-II is that choreoathetosis in the eye region during action must be interpreted with caution. The original DIS is a symmetric scale, where the same actions and rest positions are used for assessment of dystonia as well as choreoathetosis. Furthermore, all body regions are assessed in action as well as in rest. This symmetry makes the scale easy to record and score. The research group, as well as the attendees in the advisory expert meeting, considered it important to keep this symmetry in the DIS-II. This decision has consequence, however. The Rasch analysis of DIS-II could identify two misfitting items; amplitude and

duration of choreoathetosis in the eye region during action. One could then argue to delete those two items. But, since symmetry of the scale was considered important to keep, automatically it would be then impossible to assess dystonia in the eye region during action. Dystonia is known to affect participation and quality of life to a greater extent than choreoathetosis, and our research group decided to keep the two misfitting items to be able to assess dystonia in the eye region during action(162).

A possible disadvantage with the DIS-II is that impairment as well as activity limitation is included in the assessment. In the original DIS, actions the individual is not able to perform are deleted from the assessment, and accordingly the outcome measure, is presented as a percentage score of performed test items. This enables assessing the presence and severity of dystonia and choreoathetosis (impairment), without the impact these movement disorders have on function (activity limitation). In the DIS-II, the rating scale for action items have a score of "3" if the action is impossible to perform, and hence activity limitation is included in the assessment. Although complicated, it is possible to report impairment only using the DIS-II. A percentage score based on performed actions and rest postures only can be calculated in the scoring sheet, as is done in the original DIS.

Finally, although the DIS-II is considerably shorter than the original DIS, it is still comprehensive. During the development of the DIS-II, the research group discussed deleting the evaluation of the amplitude factor, and only keeping the duration factor. A valid argument for that choice would be that it is known that in individuals with dyskinetic cerebral palsy, there is a high correlation between the amplitude and duration of the dystonia and choreoathetosis(162). If we had decided to delete all amplitude items, the number of items to score would be 44 instead of 88. Despite this, we chose to keep both the amplitude and the duration factor, since the phenomenology of dystonia as well as choreoathetosis is very much characterized by the magnitude of the movements/postures as well as the amount of time that it is present during rest and action. However, if it turns out that the DIS-II is considered too comprehensive for a clinical setting, it is possible in future studies to examine construct validity and internal consistency for the duration factor only.

4.4.4 Strengths and limitations

A strength of this study was that two strategies were used to develop DIS-II, an advisory expert meeting and a Rasch measurement model analysis. Our sample size might be a limitation, since it did not allow either Differential Item Functioning analysis nor Differential Person Functioning analysis; however, due to the rareness of these disorders we considered it sufficient to analyze the rating scale of functioning, point-measure correlations, fit statistics, principal component analysis, targeting, and internal consistency.

The most important limitation of this study is probably that we did not apply the DIS-II to any patients. The aim of this study was to develop a shorter and more clinically useful assessment scale for evaluation of dystonia and choreoathetosis in children and young adults. However, we did not score any patients with the new DIS-II. Accordingly, we have not examined whether the DIS-II significantly reduces the administration and scoring time. It seems very likely that the DIS-II is less time-consuming, however, since the number of videos to record are reduced

from 26 to 14, and items to score from 144 to 88. Additionally, the rating scale is simplified, from a 5-point rating scale to a 4-point rating scale for action items, and a 3-point rating scale for rest items. According to my own experience, the original DIS takes about one hour, on average, to administer, and additionally two hours to score. I believe the DIS-II will take considerably less time to administer, and hopefully about 45 minutes to one hour to score, but that remains to be seen.

Another important limitation of our study design is that it did not allow us to change or add new test items to the DIS-II. When evaluating an existing rating scale with a Rasch measurement model analysis, using data collected from individuals assessed with the rating scale of interest, you are able to discover redundant test items. These redundant items may be deleted from the rating scale. However, it is not possible to change test items or to add new test items. Adding or changing test items is only possible if you apply your new rating scale to the population of interest, rescore the individuals with the new test instrument, and perform a second Rasch analysis.

In this present study, we used scores from 123 individuals that had been filmed according to the DIS video protocol. The aim of the study was to use those scores to evaluate construct validity and reliability (internal consistency) of the DIS-II. Our aim was not to re-score each participant with the DIS-II, and accordingly our study design did not allow us to change or delete test items. One consequence of that being that the body region trunk was deleted from the assessment scale. The reason behind this decision, as has been mentioned earlier, was that many participants were unable to perform the trunk actions, and also because the trunk during rest was considered difficult to score. It is possible, however, that it would be easier to score the trunk in other rest postures and actions, but our study design did not allow us to test that.

A possible limitation of this study was that the participants in the study were rated by several different raters. It is possible that these raters interpreted the concepts of dystonia and choreoathetosis differently, or that they used the rating scale of the DIS slightly inconsistently. One rater might be ‘more generous’ with higher scores compared to another rater. This might impact the fit statistics of the Rasch analysis, but the direction is impossible to know. A many-facets Rasch model is a method to identify and quantify ‘rater severity’ problems(163). With a many-facets Rasch model a researcher may receive information regarding item difficulty, person ability, and rater severity (and rater patterns). This can guide a test developer to change a test instrument to make it more robust and easier to score, to improve inter-rater reliability, and identify raters who deviate from the pattern and perhaps need more guidance regarding the test instrument. An important prerequisite for using a many-facets Rasch model is some suitable link across raters/test items/persons(163).

In our study, each participant was rated by only one rater, and therefore a many-facets Rasch model is not applicable. We consider this to not be of great concern, however, since earlier studies have shown a good-to-excellent inter-rater reliability and test–retest reliability for the DIS. Furthermore, when it comes to using the DIS, in the clinic as well as in research, there is already a strong recommendation to complete a DIS instructional course before using the scale. Additionally, included in the DIS scale, there are detailed definitions of dystonia and choreoathetosis in all different body regions, in rest as well as action.

5 CONCLUSIONS AND CLINICAL IMPLICATIONS

Study I: This long-term follow-up study indicates that BPTI is a benign condition that does not lead to any neurological rest symptoms. Although many children later develop migraine or other episodic syndromes that may be associated with migraine, there is a good chance that the disorder is relatively mild. We found no known mutations in candidate genes, but the etiology of BPTI is heterogenous, and for some children the natural course of BPTI may be less benign. In the case of a severe phenotype, specifically if there is a positive family history of paroxysmal disorders other than migraine, it may be important to follow the children longitudinally and genetic testing should be considered.

Study II: This study confirms that GPi-DBS is an effective treatment option for individuals with DYT-THAP1 dystonia. A total of 14 individuals were followed for a median time of about five years. All benefitted from surgery and the maximal effect was observed after a median time of 10 months. The improvement remained stable for most individuals, and at the last follow-up there was a median dystonia reduction of 58% when assessed with the BFM-M rating scale. DBS surgery should be offered to children, as well as adults, with DYT-THAP1 dystonia, even in the case of fixed contractures. Patients should be informed that the effect on orolaryngeal dystonia is likely to be less successful compared to the effect on dystonia involving the trunk and limbs. Finally, comprehensive genetic testing is recommended for all individuals with dystonia before surgery, since the underlying molecular defect might be used to predict the outcome after DBS.

Study III: This study shows that the DIS might be a valuable tool to evaluate dystonia and choreoathetosis in children and young adults with inherited or idiopathic dystonia. The DIS was developed to measure dystonia and choreoathetosis in children and young adults with dyskinetic cerebral palsy. In Study III, inter-rater reliability for the DIS, the DIS-D, and the DIS-CA was moderate-to-good, and test–retest reliability was good-to-excellent respectively, in a population of children and youth with inherited or idiopathic dystonia. Furthermore, the concurrent validity of the DIS-D was good when compared with the gold standard BFM-M. Altogether, these results indicate that the DIS might be used in inherited or idiopathic dystonia, in clinical practice as well as in research.

Study IV: In this study, the DIS-II was developed, and using a Rasch measurement model analysis, construct validity and reliability were examined. The DIS-II, a shorter and simplified version of the original DIS, evaluates dystonia and choreoathetosis in children and young adults. 11 body regions (instead of 12) are assessed during one rest posture and one action (instead of two), on a 3-step scoring scale for rest items and a 4-step scoring scale for action items respectively (instead of 5-step rating scale for all items). The outcome measure is presented as a logit measure (interval scale), which makes comparisons between individuals or over time easy. The Rasch analysis showed evidence of good construct validity and an excellent person reliability index (strata), indicating that the DIS-II may be sensitive to change. Altogether, these results support the use of the DIS-II in the clinic as well as in research.

6 FUTURE PERSPECTIVES

Study I: There are different opinions as to whether BPTI should be referred to as a benign condition or not. Delayed motor development and cognitive dysfunction have been described in several reports, but the findings are inconclusive. It is possible that there are different etiologies behind BPTI, that explain these inconsistent results. To further explore the long-term risk of additional disorders or difficulties, a case-control study could be performed, where individuals who had an earlier diagnosis of BPTI, were compared to, for example, age- and sex-matched controls. In addition, individuals with an earlier diagnosis of BPTI could be assessed using age-standardized tests (where normative data is available) evaluating cognitive function.

Study II: There is sufficient data to support the use of DBS in individuals with DYT-THAP1 dystonia. However, dystonia in the orolaryngeal region does not have a robust response to the treatment. There is some evidence that targets other than GPi, such as VIM, might be a better option to treat orolaryngeal dystonia. A suitable study design for this rare population is not straight forward. Now, in the era of WES and WGS, more individuals with DYT-THAP1 dystonia are diagnosed and, consequently, selected for DBS treatment. Additionally, international networks dealing with rare disorders have evolved, making collaboration easier. Altogether, these circumstances might enable the performance of a randomized controlled study. Otherwise, a well-designed prospective case series would, of course, be an alternative way of examining the outcome after VIM-DBS in DYT-THAP1 dystonia. Outcome measures other than the BFM are warranted, such as assessments for participation or quality of life.

Study III: The DIS was developed to evaluate dystonia and choreoathetosis in children and young adults with dyskinetic CP. We have been able to show that the DIS also provides valid and reliable outcome measures in inherited and idiopathic dystonia. Future research should focus on responsiveness. Furthermore, content validity of the DIS was considered during its development for the dyskinetic CP population. It would be interesting to evaluate content and construct validity of the DIS in isolated dystonia. One might speculate that other actions would suit this population better, and development of a modified DIS for isolated dystonia could be a possibility. The DIS has some limitations though, first it is time-consuming and secondly, the outcome measure is relative (a percentage score) making comparison difficult. A shorter version of the DIS for clinical use, with an absolute outcome measure, is warranted.

Study IV: The DIS-II, a shorter and simplified version of the DIS, was developed in Study IV. The number of body regions and actions to record and score are reduced, and the rating scale is simplified. Furthermore, the outcome measure is absolute, and may be expressed as a raw score or a logit measure. Future studies are needed to evaluate how the new 3/4 rating scale is functioning when used, since this study was built by merging categories. The DIS-II needs to be assessed for inter-rater and test–retest reliability, concurrent validity, and sensitivity to change before it can be recommended for use in clinic and intervention studies. If a modified version of the DIS-II is created for isolated dystonia, a many-facets Rasch model analysis could be performed to analyze construct validity and internal consistency. With such a model, information about how the rater contributes to the outcome measures can be examined. The result could serve as a basis for educating therapists in how to use the DIS-II.

7 ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to everyone who has contributed to this work. Special thanks go to:

All children and families that have participated in the studies. Thank you for generously offering your time to be interviewed, physically examined, video-recorded, and for leaving blood or saliva samples for genetic analysis. You have given me valuable knowledge regarding what it is like to live with a rare, and sometimes life-long, disorder with severe motor disability, and that knowledge has had great impact on my research as well as my clinical work. The findings of this thesis would not have possible without you, and hopefully other children and young adults will benefit from it in the future.

Associate Professor Kristina Tedroff, my principal supervisor, thank you for persuading me to begin this PhD project and for guiding me through these years of hard work. You have generously invited me to your research networks in Stockholm as well as in Europe. Although you always have had a full schedule, you have been there for me answering questions, even in the middle of the night when necessary. During these years we have gotten to know each other, and I now also refer to you also as a friend.

Associate Professor Lena Krumlinde Sundholm, my co-supervisor, thank you for all our discussions regarding rating scale structure, reliability, validity, ICC, and Rasch measurement model analysis. And thank you for sharing your thoughts about important things in life, such as grandchildren and gardening.

Professor Ann Nordgren, my co-supervisor, thank you for being an inspiration in research as well as clinical work. Your knowledge is outstanding, and I am so proud that you have been a part of my PhD.

Associate Professor Kirsten Holmberg, my mentor and friend, thank you for supporting me in everything. When I see a problem, you have the solution. I am so thankful that you so generously have shared so much of your life experience with me when it comes to clinical work as well as private life.

Assistant Professor Cecilia Lidbeck, co-author and friend, I truly value the time we have spent working together, all our discussions, and how generously you have shared your deep clinical knowledge on different aspects of dystonia, choreoathetosis and spasticity. Now I look forward to new adventures with you, on horseback!

Doctoral colleague Inti Vanmechelen, co-author, without you Studies III and IV would not have been possible. I admire you for your quick thinking, and I cannot imagine someone being more skilled in Excel than you.

Britt-Marie Anderlid, Tommy Stödberg, Kristina Lagerstedt-Robinson, Eva Klackenberg Arrhenius, co-authors of Study I, thank you for all the effort you put into this project.

Dan Jacobsson, Eva Åndell, Carin Andrews, Ulrike Ryll, Jenny Hedberg Graff, Linda Ek, Emma Hjalmarsson, Annika Eriksson, past and present doctoral colleagues on Floor 7, thank you all for support and encouragement.

Ann-Christin Eliasson, Kicki Löwing, Hans Forssberg, Eva Weidenhielm Broström, researchers at the Neuropediatric unit at the Department of Women`s and Children`s Health Karolinska Institutet, I would like to express my gratitude to you for having created an inclusive research environment at Floor 7.

None-Marie Kemp and Mikael Reimeringer, at the Department Women`s and Children`s Health Karolinska Institutet, thank you for all practical support during my PhD studies.

Mikael Löfstedt, senior consultant and my boss at Sachs' Children and Youth hospital, I sincerely thank you for believing in me and allowing me to take research leave.

Alejandra Cuevas, Erik Holm, Gotte Ringqvist, Sophia Boivie, Margareta Albåge, Nathalie Palmberg and Dalia Daugeliene, my colleagues at the neurologic clinic, Sachs' Children and Youth hospital. It is always a pleasure working together with you and I am so grateful that you have been taking care of all my patients during my absence.

All my friends who have supported me during periods of hard work, I am looking forward to spending more time with you.

My mother Kerstin and father Claes, you have been the best of parents during different periods of my life; first when I was a child, then when I was a young woman, and now when I am older. Thank you for always being there and supporting me!

My brother Henrik and my sister-in-law Annica, I am so grateful to have you in my life.

Ingrid and Bengt, my mother and father-in-law, our relationship is so important for me, thank you for always having time for me.

Mattias, my beloved husband and best friend, you bring joy and adventure into my life. I love you! Ellinor, Oskar, Alma, and Erik, our four wonderful children, the world is a more beautiful place because of you!

This work has been financially supported by Norrbacka-Eugeniastiftelsen, Linnea and Josef Carlssons Stiftelse, Sällskapet Barnavård, Sällsyntafonden, SÖS kliniska forskarskola, and Sachs' Children and Youth hospital (FOUU).

8 REFERENCES

1. Om sällsynta hälsotillstånd; socialstyrelsens termbank, versionsnummer: 4.3.1 2021 [updated 2021-05-12; cited 2021.] Available from: https://termbank.socialstyrelsen.se/?TermId=826&SrcLang=sv.

2. Programme of community action on rare diseases (1999-2003) Decision No 1295/99/EC of the European Parliament and of the Council of 29 April 1999 [cited 2021.] Available from: http://ec.europa.eu/health/archive/ph_overview/previous_programme/rare_diseases/raredis_wpgm99_en.pdf

3. Genetic and Rare Diseases Information Center, GARD [cited 2021.] Available from: https://rarediseases.info.nih.gov/.

4. Rare Diseases Act of 2002. [cited 2021.] Available from: https://history.nih.gov/research/downloads/PL107-280.pdf.

5. About Rare Diseases [cited 2021.] Available from: http://www.orpha.net/consor/cgi-bin/Education_AboutRareDiseases.php?lng=EN.

6. Online Mendelian Inheritance in Man OMIM® [cited 2021.] Available from: http://omim.org/.

7. Nguengang Wakap S, Lambert DM, Olry A, Rodwell C, Gueydan C, Lanneau V, et al. Estimating cumulative point prevalence of rare diseases: analysis of the Orphanet database. European journal of human genetics : EJHG. 2020;28(2):165-73.

8. Fahn S. Classification of movement disorders. Movement disorders : official journal of the Movement Disorder Society. 2011;26(6):947-57.

9. Klein C. Movement disorders: classifications. Journal of inherited metabolic disease. 2005;28(3):425-39.

10. Singer HS, Mink J, Gilbert DL, Jankovic J. Movement Disorders in Childhood. Second ed: Elsevier Science; 2015. p. 27-35.

11. Singer HS, Mink J, Gilbert DL, Jankovic J. Basal ganglia Anatomy, Biochemistry, and Physiology; Cerebellar Anatomy, Biochemistry, Physiology, and Plasticity. Movement Disorders in Childhood: Elsevier Science; 2015. p. 4-24.

12. Mink JW. Special concerns in defining, studying, and treating dystonia in children. Movement disorders : official journal of the Movement Disorder Society. 2013;28(7):921-5.

13. Fernandez-Alvarez E, Aicardi J. Movement disorders in children. 36: Cambridge University Press; 2001. p. vii-viii.

14. Fernández-Alvarez E. Transient benign paroxysmal movement disorders in infancy. European journal of paediatric neurology : EJPN : official journal of the European Paediatric Neurology Society. 2018;22(2):230-7.

15. Sanger TD, Chen D, Fehlings DL, Hallett M, Lang AE, Mink JW, et al. Definition and classification of hyperkinetic movements in childhood. Movement disorders : official journal of the Movement Disorder Society. 2010;25(11):1538-49.

16. Rosenbaum P, Paneth N, Leviton A, Goldstein M, Bax M, Damiano D, et al. A report: the definition and classification of cerebral palsy April 2006. Developmental medicine and child neurology Supplement. 2007;109:8-14.

17. Christine C, Dolk H, Platt MJ, Colver A, Prasauskiene A, Krägeloh-Mann I. Recommendations from the SCPE collaborative group for defining and classifying cerebral palsy. Developmental medicine and child neurology Supplement. 2007;109:35-8.

18. Eggink H, Kuiper A, Peall KJ, Contarino MF, Bosch AM, Post B, et al. Rare inborn errors of metabolism with movement disorders: a case study to evaluate the impact upon quality of life and adaptive functioning. Orphanet journal of rare diseases. 2014;9:177.

19. Lin JP, Lumsden DE, Gimeno H, Kaminska M. The impact and prognosis for dystonia in childhood including dystonic cerebral palsy: a clinical and demographic tertiary cohort study. Journal of neurology, neurosurgery, and psychiatry. 2014;85(11):1239-44.

20. Bäumer T, Sajin V, Münchau A. Childhood-Onset Movement Disorders: A Clinical Series of 606 Cases. Movement disorders clinical practice. 2017;4(3):437-40.

21. Silveira-Moriyama L, Lin JP. A field guide to current advances in paediatric movement disorders. Current opinion in neurology. 2015;28(4):437-46.

22. Blackburn JS, Mink JW, Augustine EF. Pediatric movement disorders: Five new things. Neurol Clin Pract. 2012;2(4):311-8.

23. Meneret A, Gaudebout C, Riant F, Vidailhet M, Depienne C, Roze E. PRRT2 mutations and paroxysmal disorders. European journal of neurology. 2013;20(6):872-8.

24. Domingo A, Erro R, Lohmann K. Novel Dystonia Genes: Clues on Disease Mechanisms and the Complexities of High-Throughput Sequencing. Movement disorders : official journal of the Movement Disorder Society. 2016;31(4):471-7.

25. Olgiati S, Quadri M, Bonifati V. Genetics of movement disorders in the next-generation sequencing era. Movement disorders : official journal of the Movement Disorder Society. 2016;31(4):458-70.

26. Dale RC, Gorman MP, Lim M. Autoimmune encephalitis in children: clinical phenomenology, therapeutics, and emerging challenges. Current opinion in neurology. 2017;30(3):334-44.

27. Abboud H, Probasco JC, Irani S, Ances B, Benavides DR, Bradshaw M, et al. Autoimmune encephalitis: proposed best practice recommendations for diagnosis and acute management. Journal of neurology, neurosurgery, and psychiatry. 2021;92(7):757-68.

28. di Biase L, Munhoz RP. Deep brain stimulation for the treatment of hyperkinetic movement disorders. Expert review of neurotherapeutics. 2016;16(9):1067-78.

29. Hale AT, Monsour MA, Rolston JD, Naftel RP, Englot DJ. Deep brain stimulation in pediatric dystonia: a systematic review. Neurosurg Rev. 2020;43(3):873-80.

30. Sanger TD, Delgado MR, Gaebler-Spira D, Hallett M, Mink JW. Classification and definition of disorders causing hypertonia in childhood. Pediatrics. 2003;111(1):e89-97.

31. Sanger TD, Chen D, Delgado MR, Gaebler-Spira D, Hallett M, Mink JW. Definition and classification of negative motor signs in childhood. Pediatrics. 2006;118(5):2159-67.

32. Snyder CH. Paroxysmal torticollis in infancy. A possible form of labyrinthitis. Am J Dis Child. 1969;117(4):458-60.

33. Drigo P, Carli G, Laverda AM. Benign paroxysmal torticollis of infancy. Brain & development. 2000;22(3):169-72.

34. Rosman NP, Douglass LM, Sharif UM, Paolini J. The neurology of benign paroxysmal torticollis of infancy: report of 10 new cases and review of the literature. Journal of child neurology. 2009;24(2):155-60.

35. Greene KA, Lu V, Luciano MS, Qubty W, Irwin SL, Grimes B, et al. Benign paroxysmal torticollis: phenotype, natural history, and quality of life. Pediatric research. 2021;90(5):1044-51.

36. Hadjipanayis A, Efstathiou E, Neubauer D. Benign paroxysmal torticollis of infancy: An underdiagnosed condition. J Paediatr Child Health. 2015;51(7):674-8.

37. Deonna T, Martin D. Benign paroxysmal torticollis in infancy. Arch Dis Child. 1981;56(12):956-9.

38. Giffin NJ, Benton S, Goadsby PJ. Benign paroxysmal torticollis of infancy: four new cases and linkage to CACNA1A mutation. Developmental medicine and child neurology. 2002;44(7):490-3.

39. Gelfand AA. Episodic syndromes of childhood associated with migraine. Current opinion in neurology. 2018;31(3):281-5.

40. Dale RC, Gardiner A, Antony J, Houlden H. Familial PRRT2 mutation with heterogeneous paroxysmal disorders including paroxysmal torticollis and hemiplegic migraine. Developmental medicine and child neurology. 2012;54(10):958-60.

41. Cuenca-Leon E, Corominas R, Fernandez-Castillo N, Volpini V, Del Toro M, Roig M, et al. Genetic analysis of 27 Spanish patients with hemiplegic migraine, basilar-type migraine and childhood periodic syndromes. Cephalalgia : an international journal of headache. 2008;28(10):1039-47.

42. Roubertie A, Echenne B, Leydet J, Soete S, Krams B, Rivier F, et al. Benign paroxysmal tonic upgaze, benign paroxysmal torticollis, episodic ataxia and CACNA1A mutation in a family. Journal of neurology. 2008;255(10):1600-2.

43. Vila-Pueyo M, Gene GG, Flotats-Bastardes M, Elorza X, Sintas C, Valverde MA, et al. A loss-of-function CACNA1A mutation causing benign paroxysmal torticollis of infancy. European journal of paediatric neurology : EJPN : official journal of the European Paediatric Neurology Society. 2014;18(3):430-3.

44. Heron SE, Dibbens LM. Role of PRRT2 in common paroxysmal neurological disorders: a gene with remarkable pleiotropy. J Med Genet. 2013;50(3):133-9.

45. Choi KD, Choi JH. Episodic Ataxias: Clinical and Genetic Features. Journal of movement disorders. 2016;9(3):129-35.

46. Brockmann K. Episodic movement disorders: from phenotype to genotype and back. Current neurology and neuroscience reports. 2013;13(10):379.

47. Claes L, Del-Favero J, Ceulemans B, Lagae L, Van Broeckhoven C, De Jonghe P. De novo mutations in the sodium-channel gene SCN1A cause severe myoclonic epilepsy of infancy. American journal of human genetics. 2001;68(6):1327-32.

48. Fuchs T, Gavarini S, Saunders-Pullman R, Raymond D, Ehrlich ME, Bressman SB, et al. Mutations in the THAP1 gene are responsible for DYT6 primary torsion dystonia. Nature genetics. 2009;41(3):286-8.

49. Groen JL, Ritz K, Contarino MF, van de Warrenburg BP, Aramideh M, Foncke EM, et al. DYT6 dystonia: mutation screening, phenotype, and response to deep brain stimulation. Movement disorders : official journal of the Movement Disorder Society. 2010;25(14):2420-7.

50. Mohammad SS, Paget SP, Dale RC. Current therapies and therapeutic decision making for childhood-onset movement disorders. Movement disorders : official journal of the Movement Disorder Society. 2019.

51. Volkmann J, Wolters A, Kupsch A, Muller J, Kuhn AA, Schneider GH, et al. Pallidal deep brain stimulation in patients with primary generalised or segmental dystonia: 5-year follow-up of a randomised trial. The Lancet Neurology. 2012;11(12):1029-38.

52. Fox MD, Alterman RL. Brain Stimulation for Torsion Dystonia. JAMA neurology. 2015;72(6):713-9.

53. Vidailhet M, Vercueil L, Houeto JL, Krystkowiak P, Lagrange C, Yelnik J, et al. Bilateral, pallidal, deep-brain stimulation in primary generalised dystonia: a prospective 3 year follow-up study. The Lancet Neurology. 2007;6(3):223-9.

54. Bruggemann N, Kuhn A, Schneider SA, Kamm C, Wolters A, Krause P, et al. Short- and long-term outcome of chronic pallidal neurostimulation in monogenic isolated dystonia. Neurology. 2015;84(9):895-903.

55. Zittel S, Moll CK, Bruggemann N, Tadic V, Hamel W, Kasten M, et al. Clinical neuroimaging and electrophysiological assessment of three DYT6 dystonia families. Movement disorders : official journal of the Movement Disorder Society. 2010;25(14):2405-12.

56. Panov F, Tagliati M, Ozelius LJ, Fuchs T, Gologorsky Y, Cheung T, et al. Pallidal deep brain stimulation for DYT6 dystonia. Journal of neurology, neurosurgery, and psychiatry. 2012;83(2):182-7.

57. Krause P, Bruggemann N, Volzmann S, Horn A, Kupsch A, Schneider GH, et al. Long-term effect on dystonia after pallidal deep brain stimulation (DBS) in three members of a family with a THAP1 mutation. Journal of neurology. 2015;262(12):2739-44.

58. Oppenheim H. Über eine eigenartige Krampfkrankheit des kindlichen und jugendlichen Alters (Dysbasia lordotica progressiva, Dystonia musculorum deformans). Neurologisches Centralblatt. 1911;30:1090-107.

59. Ozelius LJ, Hewett J, Kramer P, Bressman SB, Shalish C, de Leon D, et al. Fine localization of the torsion dystonia gene (DYT1) on human chromosome 9q34: YAC map and linkage disequilibrium. Genome research. 1997;7(5):483-94.

60. Albanese A, Bhatia K, Bressman SB, Delong MR, Fahn S, Fung VS, et al. Phenomenology and classification of dystonia: a consensus update. Movement disorders : official journal of the Movement Disorder Society. 2013;28(7):863-73.

61. Liow NY, Gimeno H, Lumsden DE, Marianczak J, Kaminska M, Tomlin S, et al. Gabapentin can significantly improve dystonia severity and quality of life in children. European journal of paediatric neurology : EJPN : official journal of the European Paediatric Neurology Society. 2015.

62. Eggink H, Coenen MA, de Jong R, Toonen RF, Eissens MH, Veenstra WS, et al. Motor and non-motor determinants of health-related quality of life in young dystonia patients. Parkinsonism & related disorders. 2019;58:50-5.

63. Fernández-Alvarez E. Prevalence of paediatric movement disorders. Pediatric movement disorders John Libbey Eurotext, UK; 2005. p. 1-18.

64. Balash Y, Giladi N. Efficacy of pharmacological treatment of dystonia: evidence-based review including meta-analysis of the effect of botulinum toxin and other cure options. European journal of neurology. 2004;11(6):361-70.

65. Albanese A, Barnes MP, Bhatia KP, Fernandez-Alvarez E, Filippini G, Gasser T, et al. A systematic review on the diagnosis and treatment of primary (idiopathic) dystonia and dystonia plus syndromes: report of an EFNS/MDS-ES Task Force. European journal of neurology. 2006;13(5):433-44.

66. Jinnah HA, Factor SA. Diagnosis and treatment of dystonia. Neurologic clinics. 2015;33(1):77-100.

67. Roubertie A, Mariani LL, Fernandez-Alvarez E, Doummar D, Roze E. Treatment for dystonia in childhood. European journal of neurology. 2012;19(10):1292-9.

68. Koy A, Lin JP, Sanger TD, Marks WA, Mink JW, Timmermann L. Advances in management of movement disorders in children. The Lancet Neurology. 2016;15(7):719-35.

69. Delnooz CC, Horstink MW, Tijssen MA, van de Warrenburg BP. Paramedical treatment in primary dystonia: a systematic review. Movement disorders : official journal of the Movement Disorder Society. 2009;24(15):2187-98.

70. Bertucco M, Sanger TD. Current and emerging strategies for treatment of childhood dystonia. Journal of hand therapy : official journal of the American Society of Hand Therapists. 2015;28(2):185-93; quiz 94.

71. Jinnah HA, Albanese A, Bhatia KP, Cardoso F, Da Prat G, de Koning TJ, et al. Treatable inherited rare movement disorders. Movement disorders : official journal of the Movement Disorder Society. 2017.

72. Fehlings D, Brown L, Harvey A, Himmelmann K, Lin JP, Macintosh A, et al. Pharmacological and neurosurgical interventions for managing dystonia in cerebral palsy: a systematic review. Developmental medicine and child neurology. 2018;60(4):356-66.

73. Wakusawa K, Haginoya K, Kitamura T, Togashi N, Ishitobi M, Yokoyama H, et al. Effective treatment with levodopa and carbidopa for hypomyelination with atrophy of the basal ganglia and cerebellum. The Tohoku journal of experimental medicine. 2006;209(2):163-7.

74. Burke RE, Fahn S, Marsden CD. Torsion dystonia: a double-blind, prospective trial of high-dosage trihexyphenidyl. Neurology. 1986;36(2):160-4.

75. Fasano A, Lozano AM. Deep brain stimulation for movement disorders: 2015 and beyond. Current opinion in neurology. 2015;28(4):423-36.

76. Moro E, LeReun C, Krauss JK, Albanese A, Lin JP, Walleser Autiero S, et al. Efficacy of pallidal stimulation in isolated dystonia: a systematic review and meta-analysis. European journal of neurology. 2017;24(4):552-60.

77. Cif L, Coubes P. Historical developments in children's deep brain stimulation. European Journal of Paediatric Neurology.21(1):109-17.

78. Reese R, Volkmann J. Deep Brain Stimulation for the Dystonias: Evidence, Knowledge Gaps, and Practical Considerations. Movement disorders clinical practice. 2017;4(4):486-94.

79. Herrington TM, Cheng JJ, Eskandar EN. Mechanisms of deep brain stimulation. Journal of neurophysiology. 2016;115(1):19-38.

80. Andrews C, Aviles-Olmos I, Hariz M, Foltynie T. Which patients with dystonia benefit from deep brain stimulation? A metaregression of individual patient outcomes. Journal of neurology, neurosurgery, and psychiatry. 2010;81(12):1383-9.

81. Elkaim LM, Alotaibi NM, Sigal A, Alotaibi HM, Lipsman N, Kalia SK, et al. Deep brain stimulation for pediatric dystonia: a meta-analysis with individual participant data. Developmental medicine and child neurology. 2019;61(1):49-56.

82. Albanese A, Sorbo FD, Comella C, Jinnah HA, Mink JW, Post B, et al. Dystonia rating scales: critique and recommendations. Movement disorders : official journal of the Movement Disorder Society. 2013;28(7):874-83.

83. Stewart K, Harvey A, Johnston LM. A systematic review of scales to measure dystonia and choreoathetosis in children with dyskinetic cerebral palsy. Developmental medicine and child neurology. 2017;59(8):786-95.

84. Pavone L, Burton J, Gaebler-Spira D. Dystonia in childhood: clinical and objective measures and functional implications. Journal of child neurology. 2013;28(3):340-50.

85. Comella CL, Leurgans S, Wuu J, Stebbins GT, Chmura T, Dystonia Study G. Rating scales for dystonia: a multicenter assessment. Movement disorders : official journal of the Movement Disorder Society. 2003;18(3):303-12.

86. Burke RE, Fahn S, Marsden CD, Bressman SB, Moskowitz C, Friedman J. Validity and reliability of a rating scale for the primary torsion dystonias. Neurology. 1985;35(1):73-.

87. Barry MJ, VanSwearingen JM, Albright AL. Reliability and responsiveness of the Barry-Albright Dystonia Scale. Developmental medicine and child neurology. 1999;41(6):404-11.

88. Monbaliu E, Ortibus E, De Cat J, Dan B, Heyrman L, Prinzie P, et al. The Dyskinesia Impairment Scale: a new instrument to measure dystonia and choreoathetosis in dyskinetic cerebral palsy. Developmental medicine and child neurology. 2012;54(3):278-83.

89. Battini R, Sgandurra G, Petacchi E, Guzzetta A, Di Pietro R, Giannini MT, et al. Movement disorder-childhood rating scale: reliability and validity. Pediatric neurology. 2008;39(4):259-65.

90. Battini R, Guzzetta A, Sgandurra G, Di Pietro R, Petacchi E, Mercuri E, et al. Scale for evaluation of movement disorders in the first three years of life. Pediatric neurology. 2009;40(4):258-64.

91. Monbaliu E, Ortibus E, Roelens F, Desloovere K, Deklerck J, Prinzie P, et al. Rating scales for dystonia in cerebral palsy: reliability and validity. Developmental medicine and child neurology. 2010;52(6):570-5.

92. Portney L, MP W. Foundations of clinical research: applications to practice. 892. Third ed: Pearson/Prentice Hall Upper Saddle River, NJ; 2009. p. 77-90.

93. Streiner DL, Norman GR, Cairney J. Health Measurement Scales: A Practical Guide to Their Development and Use. Fifth ed: Oxford University Press; 2015. p. 159-72.

94. Koo TK, Li MY. A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. Journal of chiropractic medicine. 2016;15(2):155-63.

95. Portney L, Watkins M. Foundations of clinical research: applications to practice. 892. Third ed: Pearson/Prentice Hall Upper Saddle River, NJ; 2009. p. 97-109.

96. Streiner DL, Norman GR, Cairney J. Health Measurement Scales: A Practical Guide to Their Development and Use. Fifth ed: Oxford University Press; 2015. p. 227-44.

97. Streiner DL, Norman GR, Cairney J. Health Measurement Scales: A Practical Guide to Their Development and Use. Fifth ed: Oxford University Press; 2015. p. 273-5.

98. Shaw F. Descriptive IRT vs. Prescriptive Rasch. Rasch Measurement Transactions. 5:11991. p. 131.

99. Bond T G FCM. Applying The Rasch Model, Fundamental Measurement in the Human Sciences. Second ed: New Jersey: Lawrence Erlbaum Associates; 2007. p. 29-48.

100. Bujang MA, Baharum N. A simplified guide to determination of sample size requirements for estimating the value of intraclass correlation coefficient: a review. Arch Orofac Sci. 2017;12(1):1-11.

101. Walter SD, Eliasziw M, Donner A. Sample size and optimal designs for reliability studies. Statistics in medicine. 1998;17(1):101-10.

102. Linacre J. Sample Size and Item Calibration Stability. Rasch Measurement Transactions. 1994;7:328.

103. Read A, Donnai D. New Clinical Genetics: A guide to genomic medicine. Fourth ed: Scion Publishing; 2020. p. 122-4

104. Strachan T, Read A. Human Molecular Genetics. CRC Press; 2018. p. 183-5.

105. Read A, Donnai D. New Clinical Genetics: A guide to genomic medicine. Fourth ed: Scion Publishing; 2020. p. 125-7

106. Strachan T, Read A. Human Molecular Genetics. CRC Press; 2018. p. 186-99.

107. Folio M, Fewell R. Peabody Developmental Motor Scales: Examiner's Manual and Guide to Item Administration. Austin, Tex: PRO-ED; 2000.

108. Henderson SE, Sugden DA, Barnett AL, Smits-Engelsman C. Movement assessment battery for children: Psychological Corporation London; 1992.

109. Kirkwood BR, Sterne JAC. Essential Medical Statistics. Second ed: Blackwell Science Ltd; 2003. p. 33-41.

110. Kirkwood BR, Sterne JAC. Essential Medical Statistics. Second ed: Blackwell Science Ltd; 2003. p. 344-7.

111. Portney Lg, Watkins MP. Foundations of clinical research: applications to practice. 892. Third ed: Pearson/Prentice Hall Upper Saddle River, NJ; 2009. p. 422-3.

112. Streiner DL, Norman GR, Cairney J. Health Measurement Scales: A Practical Guide to Their Development and Use. Fifth ed: Oxford University Press; 2015. p. 258-60.

113. Rosenthal R, Cooper H, Hedges L. Parametric measures of effect size. The handbook of research synthesis. 1994;621(2):231-44.

114. Kirkwood BR, Sterne JAC. Essential Medical Statistics. Second ed: Blackwell Science Ltd; 2003. p. 349-50.

115. Portney LG, Watkins MP. Foundations of clinical research: applications to practice. 892. Third ed: Pearson/Prentice Hall Upper Saddle River, NJ; 2009. p. 588-98.

116. Portney LG, Watkins MP. Foundations of clinical research: applications to practice. 892. Third ed: Pearson/Prentice Hall Upper Saddle River, NJ; 2009. p. 608-9.

117. Streiner DL, Norman GR, Cairney J. Health Measurement Scales: A Practical Guide to Their Development and Use. Fifth ed: Oxford University Press; 2015. p. 180-2.

118. Portney LG, Watkins MP. Foundations of clinical research: applications to practice. 892. Third ed: Pearson/Prentice Hall Upper Saddle River, NJ; 2009. p. 645-6.

119. Rasch Research Papers, Explorations & Explanations [May 03, 2022]. Available from: www.rasch.org.

120. JM. L. A user-guide to WINSTEPS: Grouped rating-scale model 2009 [Available from: https://www.winsteps.com/winman/isgroups.htm.]

121. Bond T FC. Applying The Rasch Model: Fundamental Measurement in the Human Sciences Second ed: New Jersey: Lawrence Erlbaum Associates; 2007. p. 222-6.

122. Linacre JM. Optimizing rating scale category effectiveness. Journal of applied measurement. 2002;3(1):85-106.

123. Linacre JM. A users guide to WINSTEPS. Rasch-model computer programs: Point-measure correlations Chicago (IL) 2009 [Available from: https://www.winsteps.com/winman/correlations.htm.]

124. Bond T FC. Applying the Rasch Model: Fundamental Measurement in the Human Sciences. Second ed: New Jersey: Lawrence Erlbaum Associates; 2007. p. 238-43.

125. Bond T FC. Applying The Rasch Model: Fundamental Measurement in the Human Sciences. Second ed: New Jersey: Lawrence Erlbaum Associates; 2007. p. 252-60.

126. WP. FJ. Rating Scale Instrument Quality Criteria. Rasch Measurement Transaction 2007 [updated 2 Jan 2018. Available from: www.rasch.org/rmt/rmt211a.htm.]

127. Bond T FC. Applying the Rasch Model: Fundamental Measurement in the Human Sciences. Second ed: New Jersey: Lawrence Erlbaum Associates; 2007. p. 40-1.

128. Wright BD MG. Number of Person or Item Strata Rasch Measurement Transactions, 2002, 16:3 p.8882002 [Available from: https://rasch.org/rmt/rmt163f.htm.]

129. The International Classification of Headache Disorders, 3rd edition (beta version). Cephalalgia : an international journal of headache. 2013;33(9):629-808.

130. Headache Classification Committee of the International Headache Society (IHS) The International Classification of Headache Disorders, 3rd edition. Cephalalgia : an international journal of headache. 2018;38(1):1-211.

131. Raieli V, Raimondo D, Cammalleri R, Camarda R. Migraine headaches in adolescents: a student population-based study in Monreale. Cephalalgia : an international journal of headache. 1995;15(1):5-12; discussion 4.

132. Ozge A, Sasmaz T, Bugdayci R, Cakmak SE, Kurt AO, Kaleagasi SH, et al. The prevalence of chronic and episodic migraine in children and adolescents. European journal of neurology. 2013;20(1):95-101.

133. Marini C, Conti V, Mei D, Battaglia D, Lettori D, Losito E, et al. PRRT2 mutations in familial infantile seizures, paroxysmal dyskinesia, and hemiplegic migraine. Neurology. 2012;79(21):2109-14.

134. Shin M, Douglass LM, Milunsky JM, Rosman NP. The Genetics of Benign Paroxysmal Torticollis of Infancy: Is There an Association With Mutations in the CACNA1A Gene? Journal of child neurology. 2016;31(8):1057-61.

135. Humbertclaude V, Krams B, Nogue E, Nagot N, Annequin D, Tourniaire B, et al. Benign paroxysmal torticollis, benign paroxysmal vertigo, and benign tonic upward gaze are not benign disorders. Developmental medicine and child neurology. 2018;60(12):1256-63.

136. Indelicato E, Nachbauer W, Karner E, Eigentler A, Wagner M, Unterberger I, et al. The neuropsychiatric phenotype in CACNA1A mutations: a retrospective single center study and review of the literature. European journal of neurology. 2019;26(1):66-e7.

137. Humbertclaude V, Riant F, Krams B, Zimmermann V, Nagot N, Annequin D, et al. Cognitive impairment in children with CACNA1A mutations. Developmental medicine and child neurology. 2020;62(3):330-7.

138. Jinnah HA, Alterman R, Klein C, Krauss JK, Moro E, Vidailhet M, et al. Deep brain stimulation for dystonia: a novel perspective on the value of genetic testing. Journal of neural transmission (Vienna, Austria : 1996). 2017;124(4):417-30.

139. Artusi CA, Dwivedi A, Romagnolo A, Bortolani S, Marsili L, Imbalzano G, et al. Differential response to pallidal deep brain stimulation among monogenic dystonias: systematic review and meta-analysis. Journal of neurology, neurosurgery, and psychiatry. 2020;91(4):426-33.

140. Tisch S, Kumar KR. Pallidal Deep Brain Stimulation for Monogenic Dystonia: The Effect of Gene on Outcome. Frontiers in neurology. 2020;11:630391.

141. Cif L, Vasques X, Gonzalez V, Ravel P, Biolsi B, Collod-Beroud G, et al. Long-term follow-up of DYT1 dystonia patients treated by deep brain stimulation: an open-label study. Movement disorders : official journal of the Movement Disorder Society. 2010;25(3):289-99.

142. Panov F, Gologorsky Y, Connors G, Tagliati M, Miravite J, Alterman RL. Deep brain stimulation in DYT1 dystonia: a 10-year experience. Neurosurgery. 2013;73(1):86-93; discussion

143. Meyer E, Carss KJ, Rankin J, Nichols JM, Grozeva D, Joseph AP, et al. Mutations in the histone methyltransferase gene KMT2B cause complex early-onset dystonia. Nature genetics. 2017;49(2):223-37.

144. Deutschlander A, Asmus F, Gasser T, Steude U, Botzel K. Sporadic rapid-onset dystonia-parkinsonism syndrome: failure of bilateral pallidal stimulation. Movement disorders : official journal of the Movement Disorder Society. 2005;20(2):254-7.

145. Albanese A, Di Giovanni M, Amami P, Lalli S. Failure of pallidal deep brain stimulation in DYT12-ATP1A3 dystonia. Parkinsonism & related disorders. 2017;45:99-100.

146. Larsh T, Wu SW, Vadivelu S, Grant GA, O'Malley JA. Deep Brain Stimulation for Pediatric Dystonia. Seminars in pediatric neurology. 2021;38:100896.

147. Lumsden DE, Kaminska M, Ashkan K, Selway R, Lin JP. Deep brain stimulation for childhood dystonia: Is 'where' as important as in 'whom'? European journal of paediatric neurology : EJPN : official journal of the European Paediatric Neurology Society. 2017;21(1):176-84.

148. Sanger TD, Robison A, Arguelles E, Ferman D, Liker M. Case Report: Targeting for Deep Brain Stimulation Surgery Using Chronic Recording and Stimulation in an Inpatient Neuromodulation Monitoring Unit, With Implantation of Electrodes in GPi and Vim in a 7-Year-Old Child With Progressive Generalized Dystonia. Journal of child neurology. 2018;33(12):776-83.

149. Ostrem JL, San Luciano M, Dodenhoff KA, Ziman N, Markun LC, Racine CA, et al. Subthalamic nucleus deep brain stimulation in isolated dystonia: A 3-year follow-up study. Neurology. 2017;88(1):25-35.

150. Xu W, Li H, Zhang C, Sun B, Wu Y, Li D. Subthalamic Nucleus Stimulation in Pediatric Isolated Dystonia: A 10-Year Follow-up. The Canadian journal of neurological sciences Le journal canadien des sciences neurologiques. 2020;47(3):328-35.

151. Sanger TD, Liker M, Arguelles E, Deshpande R, Maskooki A, Ferman D, et al. Pediatric Deep Brain Stimulation Using Awake Recording and Stimulation for Target Selection in an Inpatient Neuromodulation Monitoring Unit. Brain sciences. 2018;8(7).

152. Poologaindran A, Ivanishvili Z, Morrison MD, Rammage LA, Sandhu MK, Polyhronopoulos NE, et al. The effect of unilateral thalamic deep brain stimulation on the vocal dysfunction in a patient with spasmodic dysphonia: interrogating cerebellar and pallidal neural circuits. Journal of neurosurgery. 2018;128(2):575-82.

153. Mure H, Morigaki R, Koizumi H, Okita S, Kawarai T, Miyamoto R, et al. Deep brain stimulation of the thalamic ventral lateral anterior nucleus for DYT6 dystonia. Stereotactic and functional neurosurgery. 2014;92(6):393-6.

154. Kundu B, Schrock L, Davis T, House PA. Thalamic Deep Brain Stimulation for Essential Tremor Also Reduces Voice Tremor. Neuromodulation : journal of the International Neuromodulation Society. 2018;21(8):748-54.

155. Badhiwala JH, Karmur B, Elkaim LM, Alotaibi NM, Morgan BR, Lipsman N, et al. Clinical phenotypes associated with outcomes following deep brain stimulation for childhood dystonia. Journal of neurosurgery Pediatrics. 2019:1-9.

156. Vasques X, Cif L, Gonzalez V, Nicholson C, Coubes P. Factors predicting improvement in primary generalized dystonia treated by pallidal deep brain stimulation. Movement disorders : official journal of the Movement Disorder Society. 2009;24(6):846-53.

157. Visser JE, Cotton AC, Schretlen DJ, Bloch J, Tedroff K, Schechtmann G, et al. Deep brain stimulation in Lesch-Nyhan disease: outcomes from the patient's perspective. Developmental medicine and child neurology. 2021;63(8):963-8.

158. Monbaliu E, Ortibus E, Prinzie P, Dan B, De Cat J, De Cock P, et al. Can the Dyskinesia Impairment Scale be used by inexperienced raters? A reliability study. European journal of paediatric neurology : EJPN : official journal of the European Paediatric Neurology Society. 2013;17(3):238-47.

159. Vanmechelen I, Dan B, Feys H, Monbaliu E. Test-retest reliability of the Dyskinesia Impairment Scale: measuring dystonia and choreoathetosis in dyskinetic cerebral palsy. Developmental medicine and child neurology. 2020;62(4):489-93.

160. World Health Organization G. International Classification of Functioning, Disability and Health. 2001.

161. Danielsson A, Vanmechelen I, Lidbeck C, Krumlinde-Sundholm L, Ortibus E, Monbaliu E, et al. Reliability and Validity of the Dyskinesia Impairment Scale in Children

and Young Adults with Inherited or Idiopathic Dystonia. Journal of clinical medicine. 2020;9(8).

162. Monbaliu E, de Cock P, Ortibus E, Heyrman L, Klingels K, Feys H. Clinical patterns of dystonia and choreoathetosis in participants with dyskinetic cerebral palsy. Developmental medicine and child neurology. 2016;58(2):138-44.

163. Bond T G FCM. Applying The Rasch Model, Fundamental Measurement in the Human Sciences. Second ed: New Jersey: Lawrence Erlbaum Associates; 2017. p. 143-59.