Doctoral theses at NTNU, 2016:129

Ann-Kristin Gunnes Elvrum

Assessment of hand function in children with bilateral cerebral palsy

Development and measurement properties of outcome measures and classifications

NTNU
Norwegian University of Science and Technology

Doctoral thesis

Ann-Kristin Gunnes Elvrum

Doctoral theses at NTNU, 2016:129

NTNU
Norwegian University of Science and Technology
Thesis for the Degree of Philosophiae Doctor
Faculty of Medicine
Department of Laboratory Medicine, Children's and Women's Health

ISBN 978-82-326-1596-4 (printed ver.)
ISBN 978-82-326-1597-1 (electronic ver.)
ISSN 1503-8181

NTNU
Norwegian University of Science and Technology

NTNU

Ann-Kristin Gunnes Elvrum

Assessment of hand function in children with bilateral cerebral palsy

Development and measurement properties of outcome measures and classifications

Thesis for the Degree of Philosophiae Doctor

Trondheim, May 2016

Norwegian University of Science and Technology
Faculty of Medicine
Department of Laboratory Medicine,
Children's and Women's Health

NTNU
Norwegian University of Science and Technology

NTNU
Norwegian University of Science and Technology

Thesis for the Degree of Philosophiae Doctor

Faculty of Medicine
Department of Laboratory Medicine, Children's and Women's Health

© Ann-Kristin Gunnes Elvrum

ISBN 978-82-326-1596-4 (printed ver.)
ISBN 978-82-326-1597-1 (electronic ver.)
ISSN 1503-8181

Doctoral theses at NTNU, 2016:129

Printed by NTNU Grafisk senter

SAMMENDRAG

Mange barn med cerebral parese har vansker med å utføre daglige aktiviteter i hjem, skole eller barnehage på grunn av nedsatt håndfunksjon Til tross for det store omfanget av slike vansker, så er behandlingen som tilbys for å bedre håndfunksjonen i liten grad kunnskapsbasert. Dette gjelder spesielt behandling for å bedre håndfunksjonen hos barn med nedsatt funksjon i begge hender (bilateral CP). Hovedårsaken til dette er at man mangler gyldige og pålitelige måleredskaper som er følsomme nok til å måle endring. Funksjonelle klassifikasjoner og standardiserte tester med gode måleegenskaper er avgjørende for å kunne følge opp barnets naturlige utvikling og gi råd rundt forventet utvikling, samt for å kunne planlegge og evaluere aktuelle behandlingstilbud.

De viktigste målene med denne avhandlingen er derfor (1) å undersøke gyldighet (validitet) og pålitelighet (reliabilitet) til en mye brukt klassifisering av finmotorikk hos barn med CP, og (2) å bidra til bedre planlegging og evaluering av aktuelle tiltak ved å identifisere og utvikle standardiserte håndfunksjonstester med gode måleegenskaper for barn med bilateral CP.

Studie I og II undersøker validitet og reliabilitet av Bimanual Fine Motor Function (BFMF), som er en klassifisering av finmotorisk funksjon hos barn med CP. I studie I besto datasettet av BFMF og Manual Ability Classification System (MACS) klassifiseringer fra 539 barn med CP, mens i studie II ble 79 barn med CP inkludert. Studie III og IV undersøker måleegenskapene til standardiserte tester som evaluerer håndfunksjon hos barn med bilateral CP. Studie III er en systematisk litteraturgjennomgang som identifiserer tilgjengelige håndfunksjonstester og vurderer måleegenskapene til disse testene, mens studie IV beskriver utviklingen og valideringen av den nye testen, Both Hands Assessment (BoHA), der data fra 171 barn med bilateral CP ble benyttet.

Resultatene indikerer at BFMF er en valid og reliabel klassifisering av finmotorisk kapasitet hos barn med CP. Ved bruk av både BFMF og MACS, som klassifiserer faktisk bruk av hendene, kan man få en mer utfyllende beskrivelse av barnets håndfunksjon. Gjennom den systematiske litteraturgjennomgangen identifiserte vi et foreldre-rapportert spørreskjema og fire håndfunksjonstester. Spørreskjemaet måler hvor lett eller vanskelig det er for barnet å utføre ulike daglige aktiviteter som krever bruk av hendene, men gir ingen informasjon om det er en eller begge hender som brukes. Tre av håndfunksjonstestene evaluerer kapasitet til å bruke hver hånd for seg, mens den fjerde hovedsakelig evaluerer om barnet har aldersadekvat funksjon i dominant hånd. Best evidens for validitet og reliabilitet ble funnet for

spørreskjemaet ABILHAND-kids og for testen Melbourne Assessment 2, men informasjon om følsomhet for endring mangler for begge disse instrumentene. Vi identifiserte ingen håndfunksjonstester som på en valid og reliabel måte måler hvordan barn med bilateral CP håndterer gjenstander ved bruk av begge hender. Derfor utviklet og validerte vi en ny test, Both Hands Assessment (BoHA), for å måle tohåndsfunksjon hos barn med bilateral CP (MACS nivå I-III). Resultatet av dette viser at BoHA er en valid test for barn med bilateral CP når den deles inn i to versjoner; en for barn med asymmetrisk (BoHA-A) og en for barn med mer symmetrisk bruk av hendene (BoHA-S). Ankring av disse to versjonene til samme måleskala gjør at resultatene som oppnås kan sammenlignes uavhengig av om barnet har en asymmetrisk eller mer symmetrisk bruk av hendene.

Denne avhandlingen viser at BFMF er en valid og reliabel klassifisering av finmotorisk funksjon hos barn med CP. Bruk av BFMF sammen med MACS vil kunne gi en mer nyansert beskrivelse av håndfunksjonen hos barn med CP, noe som er nyttig både i klinikk og forskning. For å kunne planlegge å måle effekt av ulike tiltak er det imidlertid behov for valide og reliable tester som er følsomme for endring. Melbourne Assessment 2 kan benyttes for å måle enhåndskapasitet hos barn med bilateral CP, mens ABILHAND-kids kan benyttes for å måle hvor lett eller vanskelig det er for barnet å utføre daglige aktiviteter som krever bruk av hendene. Det er imidlertid usikkert hvor godt disse måleinstrumentene fanger opp endring. Den nye testen BoHA gir en valid måling av spontan bruk av begge hender hos barn med bilateral CP og tilbyr dermed et nytt perspektiv som kan være nyttig ved planlegging og evaluering av tiltak. Fremtidige studier som undersøker reliabilitet og hvor godt BoHA fanger opp endring er imidlertid nødvendige.

Kandidat: Ann-Kristin Gunnes Elvrum
Institutt: Institutt for laboratoriemedisin, barne- og kvinnesykdommer
Veiledere: Torstein Vik, Guro L. Andersen, Lena Krumlinde-Sundholm og Jon Skranes
Finansieringskilder: Klinikk for kliniske Servicefunksjoner, St. Olavs Hospital, Stiftelsen Frimurare Barnhuset i Stocholm, Samarbeidsorganet mellom Helse Midt-Norge og NTNU

Overnevnte avhandling er funnet verdig til å forsvares offentlig
for graden PhD i klinisk medisin.
Disputas finner sted i auditoriet i MTFS
torsdag 19. mai, klokken 12.15

The journey of a thousand miles begins with one step.

Lao Tzu

Start by doing what's necessary; then do what's possible; and suddenly you are doing the impossible.

Francis of Assisi

Contents

ACKNOWLEDGEMENTS

This work was carried out at the Department of Laboratory Medicine, Children's and Woman's Health at the Norwegian University of Sciences and Technology (NTNU), founded through the Department of Clinical Services at the St. Olavs Hospital, Stiftelsen Frimurare Barnhuset in Stocholm, and Samarbeidsorganet between the Central Norway Regional Health Authority and NTNU.

In addition to providing valuable financial support, the Department of Clinical Services at St. Olavs Hospital kindly arranged for the necessary leave of absence from my clinical position in order to work on this thesis. As a clinician to be given this opportunity to perform research has enriched my life, both in terms of increased knowledge, but also by the large amount of wonderful people that I have met and come to know. Therefore, I would like to greatly thank head of department Lise Lundbom Støylen, and immediate superior Marianne Mesteig, who believed in me and encouraged me both mentally and financially to start on this PhD thesis in 2011.

This PhD thesis would not have been possible without the support from all of you around me. First and foremost, I would like to express my sincerest gratitude to my main supervisor Torstein Vik. You have patiently shared your vast knowledge with me and have this great gift of understanding the essence of what I am trying to formulate, usually long before I do! Thank you for believing in me, encouraging me, educating me, and providing me with confidence (including giving me small "pushes" when I needed that). Also thank you for introducing me to a broad network of researchers around the world. I appreciate so much the person you are; the humour, the way you meet everyone around you and your broad interests.

Great thanks to my co-supervisor Lena Krumlinde Sundholm, for including me in the development of the new test, BoHA, together with Britt-Marie Zetrhæeus. Thank you both for sharing your extensive knowledge with me. I have learned so much through this project and have thoroughly enjoyed working with you! Not to mention, working together at beautiful Sippsö, and sailing back to Stockholm in the sunset. Thank you also for including me in the research group at Karolinska and in the AHA network. The presentations, discussions and constructive comments at the research meetings have been of great value and inspiration, as well as getting to know your wonderful colleagues.

Warm thanks also to my co-supervisor Guro L. Andersen for being a great inspirer and for your support and encouragement during ups and downs, especially with the first manuscript. This study would not have been possible without you and Sandra. Thanks also for all the fun we have had during lunch breaks, meetings around the world, and for willingly offering me a spare bed at your hotel room whenever I needed that.

I would also like to thank warmly my co-authors in the separate papers for their substantial contributions and support: Kate Himmelmann, Eva Beckung, Ann-Marie Öhrvall, Stian Lydersen, Ingrid Riphagen and Rannei Sæther. It has been a great pleasure working together with each one of you. You have all added important knowledge to my work.

Warm thanks, to former and current members of the CP research group, the CEBRA group, the Journal Club, and to colleagues and roommates at the sixth floor, for interesting discussions, encouragement and support, and most of all for making it so much fun to go to work. Special thanks to Rannei and Siri for all I have learned from you, for all the fun we have had, and for always taking the time to discuss work related and personal issues whenever I have needed that. Thanks also to Torarin, who from the very first beginning expected me to be the "hand expert", now you can see what that led to.

Great thanks go to all my dear colleagues at the occupational therapy department for your support and interest, and for laughter and fun at lunch breaks and OT trips. Special thanks go to Margunn. Thank you for being such a dear colleague and friend. The regular talks and discussions with you are necessary ingredients in both my professional and personal life. Thanks also to my colleagues at the Department of Pediatrics and the Department of Clinical Services, St. Olavs Hospital. I am looking forward to "get back in business" together with you all.

Warm thanks, to all my friends, family and extended family, for your kindness, encouragement and friendship. You all mean so much to me! Special thanks go to my parents, Margrethe and Alf, for always supporting me and believing in me, and for teaching me the value of hard work and the recreational joy of nature. Thanks to my dear sister Ruth and my brother Lars for all the things we have shared together throughout our lives. I hope you know how important you are to me. And most of all, thanks to my dearest Knut and our beloved boys, Espen, Eivind and Sindre, for your love and support, for your patience with me, and most of all for constantly reminding me of what is most important in life.

Lastly, thanks to all the children, and their parents, who participated in the studies, and to all the occupational therapists in Norway and Sweden who contributed in the data collection. Your positive attitude, patience and enthusiasm made this thesis possible.

LIST OF PAPERS

Paper I

Elvrum AKG, Andersen GL, Himmelmann K, Beckung E, Öhrvall AM, Lydersen S, Vik T. Bimanual Fine Motor Function (BFMF) classification in children with cerebral palsy: aspects of construct and content validity. *Physical & Occupational Therapy in Pediatrics*. 2016 Feb; 36 (1): 1-16. Epub 2014 Nov 6. PMID: 25374154

Paper II

Elvrum AKG, Beckung E, Sæther R, Lydersen S, Vik T, Himmelmann K. Bimanual capacity of children with cerebral palsy: Intra- and inter-rater reliability of a revised edition of the Bimanual Fine Motor Function classification. Revised version submitted.

Paper III

Elvrum AKG, Sæther R, Riphagen I, Vik T. Clinical outcome measures evaluating hand function in children with bilateral cerebral palsy: A systematic review. Revised version submitted.

Paper IV

Elvrum AKG, Zetrhæus BM, Vik T, Krumlinde-Sundholm L. Development of a new test of hand function, Both Hands Assessment (BoHA) for children with bilateral cerebral palsy. Submitted.

ABBREVIATIONS

A = Activity

AHA = Assisting Hand Assessment

BF = Body Function

BFMF = Bimanual Fine Motor Function

BFMF 2 = Bimanual Fine Motor Function second edition

BoHA = Both Hands Assessment

BoHA-A = Both Hands Assessment Asymmetry version

BoHA-S = Both Hands Assessment Symmetry version

C = Criterion

CI=Confidence Interval

COSMIN = COnsensus-based Standards for the selection of health Measurement INstruments CP = Cerebral Palsy

CPRN = Cerebral Palsy Register of Norway

D = Dominant hand

DIF = Differential Item Functioning

EDPA = Erhardt Developmental Prehension Assessment

ICC = Intraclass Correlation Coefficient

ICF = International Classification of Functioning, Disability and Health

ICF-CY = International Classification of Functioning, Disability and Health, Children and Youth

K_w = Cohen's weighted Kappa

MA = Melbourne Assessment of Unilateral Upper Limb Function

MA2 = Melbourne Assessment 2

MACS = Manual Ability Classification System

ND = Non-Dominant hand

p = Probability value

P = Performance

PCA = Principal Component Analysis

PDMS = Peabody Developmental Motor Scales

PDMS-2 = Peabody Developmental Motor Scales Second Edition

QUEST = Quality of Upper Extremity Skills Test

r = Pearson product-moment coefficient of correlation

rho= Spearman's rank correlation coefficient

SUMMARY

Manual impairments in children with cerebral palsy (CP) have been found to have significant effects on independence in daily activities. Despite these well-known consequences, there are few evidence-based interventions that have been found to be effective in improving hand function. This is especially evident in children with impaired function in both hands (bilateral CP). The main challenge is the lack of outcome measures that are sensitive to change. Standardized outcome measures with sound measurement properties are crucial for prognostic counselling, for the assessment of the development of hand function, and for appropriate planning and evaluation of interventions.

In addition to outcome measures that are able to measure change, classification scales with sound measurement properties are needed in order to document results and in order to use results obtained in intervention studies to compare and generalize to other populations.[1] Furthermore, functional classifications are used in epidemiological studies to increase our understanding of distribution and causation of functional limitations in children with CP.[2] In the common data base of CP registers established by the Surveillance of Cerebral Palsy in Europe (SCPE) the principal classification of hand function is the Bimanual Fine Motor Function (BFMF) classification. However, the information on measurement properties of this classification is scarce and it was therefore necessary to investigate validity and reliability of the BFMF to consider whether continued use of this classification could be justified.

The main aim of this doctoral work was therefore to develop and validate assessments (i.e. classifications and outcome measures) of hand function in children with bilateral CP, to enable planning and evaluation of interventions. The specific aims were (1) to investigate the validity and reliability of a much used classification of fine motor function in children with CP (papers I-II), and (2) to identify and develop standardized tests with sound measurement properties evaluating hand function in children with bilateral CP (papers III-IV).

Paper I investigated construct validity of the BFMF classification in children with CP using data from the CP Registers of Norway and western-Sweden classifying BFMF and Manual Ability Classification System (MACS) levels in 539 children with CP. In addition, the contents of the BFMF and MACS were compared using the international classification of functioning children and youth version (ICF-CY) as a common reference.[3,4] The results indicated that the BFMF is a valid classification of fine motor capacity in children with CP, giving supplementary information to the MACS which classifies actual use of the hands.

For paper II, the intra- and inter-rater reliability of the BFMF was explored. Four raters classified fine motor capacity in 79 Norwegian and Swedish children with CP and the results suggested high reliability of the BFMF. The findings from papers I and II indicate that the BFMF has appropriate measurement properties for continued use in population-based CP registers and in research to classify fine motor capacity. Furthermore, the BFMF used in the clinic may contribute to easily accessible information regarding a child's capacity to grasp and manipulate.

Paper III was a systematic literature review appraising measurement properties of outcome measures used to assess hand function in children with bilateral CP. Five hand function measures were identified where measurement properties had been evaluated in children with bilateral CP. The Melbourne Assessment 2 (MA2), measuring unimanual capacity, and the ABILHAND-Kids, assessing perceived manual ability, had the strongest level of evidence for aspects of both reliability and validity. However, further research is required to determine the responsiveness of these measures. None of the identified outcome measures evaluated actual bimanual performance (i.e. spontaneous handling of objects requiring the use of both hands) in children with bilateral CP.

Paper IV describes the development and validation of the new test, Both Hands Assessment (BoHA), including data from 171 Norwegian and Swedish children with bilateral CP. The contents of the BoHA items were generated based on observations of bimanual play in children with bilateral CP, and an adaptation of the Assisting Hand Assessment (AHA), using expert judgement. The BoHA scale items were further refined and validated by the use of Rasch measurement model analysis, allowing for the transformation of ordinal raw scores into interval scale measures. Strong evidence of internal scale validity and aspects of reliability was found for BoHA when separated into two versions: one for children with asymmetric hand use (BoHA-A) and one for children with more symmetric hand use (BoHA-S). The two versions were linked through anchoring of items, creating a common measure scale of bimanual performance while still allowing use of separate item difficulty hierarchies. Thus, the BoHA can be used to describe and compare bimanual performance in the heterogeneous group of children with bilateral CP and thereby contribute to increased knowledge regarding how these children use their hands together when handling objects. Furthermore, the BoHA has the potential to become a valuable outcome measure to guide treatment and evaluate its effect in children with bilateral CP.

1. INTRODUCTION

Effective use of the hands is essential for the performance of most daily activities. For children with cerebral palsy (CP) motor impairments in one (unilateral) or both sides of the body (bilateral) causes varying degrees of limitations in functional use of the hands.[5] Regardless of the degree of severity, decreased hand function has an impact on the children's daily activities in self-care, school, play and leisure.[6]

As an occupational therapist working in the neuro-orthopedic team of St. Olavs Hospital, my main focus was to promote independence and participation in daily activities in children with CP, through appropriate assessments and interventions targeting the children's hand function. A lot of time and resources were put into this, but a significant problem was the lack of outcome measures developed and validated for children with CP that were sensitive enough to indicate change after interventions.

A big step forward in this regard was the launching of the Assisting Hand Assessment (AHA) for children with unilateral CP in 2003.[7] The AHA is a highly valued measure both in the clinic and in research, and has been found to have strong measurement properties.[8-12] With the AHA we have a standardized test that can be used in the clinic that is playful, but at the same time provides valuable information regarding effective and spontaneous use of hands when handling toys that require the use of both hands. Two AHA versions have been developed. The Small-Kids and the School-Kids AHA are developed for children aged 18 months to 5 years, and 6 to 12 years respectively. With strong alternate-form reliability, these versions allow us to describe, understand and monitor longitudinal development of hand function in children with unilateral CP, as well as inform treatment planning and evaluate outcomes of intervention.[8,13,10] Furthermore, the use of the AHA in research has contributed to increased evidence-based knowledge regarding efficacy of interventions targeting hand function in children with unilateral CP.

For children with bilateral CP, no similar test of hand function was available and consequently evidence-based knowledge regarding efficacy of interventions was scarce. Knowing that the AHA measures the effective use of the affected hand in bimanual activities, this test was not intended to measure bimanual performance in children who had varying degrees of functional limitations in both hands. So the lack of outcome measures developed and validated for children with bilateral CP was a significant hindrance in the assessment, planning and evaluation of interventions targeting hand function in these children.

Thus, I was very motivated to contribute to the development of a new test of bimanual performance in children with bilateral CP, the Both Hands Assessment (BoHA), in collaboration with the research group who developed the AHA. The new test, BoHA, is based on the AHA, but modified and validated for children with bilateral CP. The large variation in how these children used their hands was a challenge when describing the criteria defining the various BoHA items, as well as in the validation of the test. In this process, the use of Rasch measurement model analyses was essential and resulted in two BoHA versions that were linked through anchoring of items, creating a common measure scale of bimanual performance. Hopefully the BoHA will contribute to increased knowledge of natural development of hand function in children with bilateral CP and be useful in planning and evaluation of interventions.

In addition to standardized outcome measures that are able to measure change, classification scales with appropriate measurement properties are needed to describe and group children according to common characteristics, so that populations of children can be compared and results from intervention studies can be generalized to other populations.[1] In the CP Register of Norway (CPRN) the principal classification of hand function is the Bimanual Fine Motor Function (BFMF) classification. This classification was developed in 2002, but was still not validated and information on reliability was limited. Thus, in collaboration with the researchers involved in the CP registers in Norway and western Sweden we investigated measurement properties of the BFMF and found that the BFMF is a valid and reliable classification of fine motor function, and since no other classifications describe this aspect of hand function we believe that the BFMF will continue to contribute with important information in population-based studies and probably also in the clinic.

Through the development and validation of outcome measures provided in this thesis, it is my sincere hope that hand function assessments and interventions in children with bilateral CP will be more evidence-based. Children with bilateral CP spend much of their time on treatment and training, and use great amounts of health resources. It is therefore important to provide a necessary and reliable basis for describing a child’s hand function, evaluating services and measuring clinical change, to avoid spending resources on inefficient treatment, or treatments that does not matter for the individual child.

2. BACKGROUND

In this section the unique role of the hands in the performance of daily activities, as well as important features of the International Classification of Functioning, Disability and Health (ICF) related to hand function are presented (sections 2.1-2.2). In addition, the heterogeneous diagnosis cerebral palsy (CP) will be defined according to commonly used classifications, and definitions and methods for assessing hand function will be introduced (sections 2.3-2.4). Further, aspects of hand function in children with CP relevant to the current thesis will be addressed (section 2.5).

2.1 The unique role of the hands in the performance of everyday activities

The ability to use our hands is vital to our interaction with the environment, both through non-verbal communication, human contact and through exploration and manipulation of objects. The flexibility and adaptability of the human hand is remarkable. It can separate rough from smooth or cold from hot, and can act as a multifaceted tool able to perform extremely gentle and precise actions or heavy labor.[14] Already in the 3rd century, the Greek philosopher and scientist Aristoteles stated that *"…the hand is the tool of tools…"* and *"…the hand is for the body as the intellect is for the soul."*[15] The great importance of hand skills in humans is also reflected in the large areas of the brain dedicated to hand movement (see illustrative homunculus: Figure 1). It is the control of the spinal moto neurons by the cerebral cortex that provides the hand with its notable motor repertoire.[16]

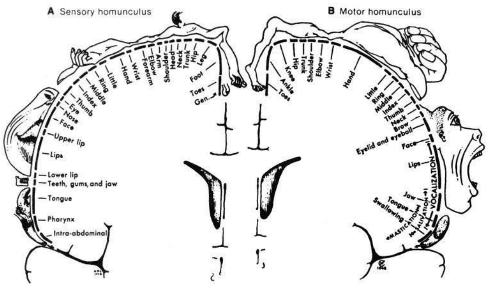


Figure 1: The illustrative sensory and motor homunculus provided by Penfield and Rasmussen in 1950.[17]

Functional performance in almost all life situations requires the handling of objects; thus the acquisition of hand skills is essential for children's development and participation in activities of daily life, such as self-care, education, play and leisure.[14,18] For skillful handling of objects, coordinated and goal-directed movements for reaching, grasping, and manipulating need to be adjusted to the different object properties (weights, shapes, textures) and to the location and orientation of the objects in the environment.[19,20] This depends on appropriate connections between many areas of the brain responsible for planning and fine-tuning of movements based on integration and perception of somatosensory and visual information.[21] Furthermore, the musculoskeletal components are crucial as the effectors of the planned movement. In addition cognitive abilities such as motivation, attention, concentration and task comprehension are important for the inner drive to explore objects, for the perfection of movements and for the development of appropriate memory strategies for smooth and coordinated movements.[20]

Most of the activities performed in daily life require the cooperative use of both hands. An important aspect of skilled hand use is therefore the ability to perform different types of collaborative actions with the hands.[20] Bimanual holding, transfer of objects from one hand to the other, pulling objects apart, pushing objects together, or manipulating objects with one hand while holding with the other require more or less symmetrical or asymmetrical actions with the hands, depending on object properties and task requirements.[22]

Moreover, the use of the hands in daily activities cannot be understood outside of context. Performance of everyday activities and development of hand skills takes place in a dynamic interaction between the child, the functional task or activity, and the environment (cultural, institutional, physical and social). Consequently, strengths or limitations within one of these areas will most certainly affect the other areas.[18] This is confirmed in studies indicating that limitations in the ability to handle objects affect the child's functioning and performance of activities in daily life.[6,20]

2.2 International Classification of Functioning, Disability and Health (ICF)

The dynamic relationships between functioning and disability in relation to health can also be described and understood by using the universal framework and common language provided by the International Classification of Functioning, Disability and Health (ICF),[23] and its version for Children and Youth (ICF-CY).[24] According to this framework, functioning comprises body functions and structures, and activity and participation, while the term disability describes impairments (i.e. problems in body function or structure), activity limitations and/or participation restrictions. In addition, the impact of contextual factors on functioning and disability can be described through the components environmental factors and personal factors (Figure 2).

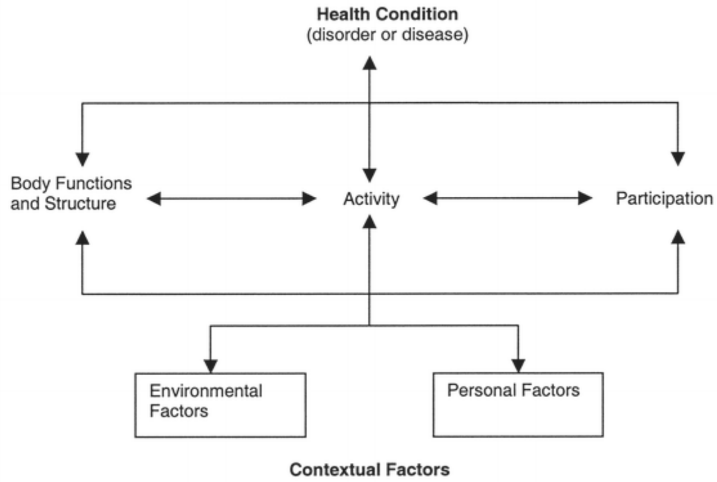


Figure 2: International Classification of Functioning, disability and health (ICF).[23]

The activity and participation component of the ICF-CY can be further considered according to the qualifiers capacity and performance. Capacity describes what the child can do at his or her best, whereas performance describes what the child usually does in the real world.[24]

Using the ICF-CY framework, the consequences of neurological impairments (body structure) can be considered according to body functions (e.g. impaired muscle tone, strength or range of motion), activity (e.g. limited hand function causing difficulties performing everyday tasks and activities), and participation (e.g. restricted involvement in life). The framework also indicates that there may be differences between the child's best capacity to use its hands under optimal circumstances and actual use of the hands in daily life (i.e. performance). This may be dependent on personal factors such as motivation and cognitive ability; the physical and social environments in which the child lives (e.g., home, community, and school environment) will also have an impact on the child's functional use of the hands and opportunities for activity and participation.[24]

2.3 Cerebral palsy

Cerebral palsy is the most common cause of severe physical disability in childhood with a prevalence of around 2 per 1000 live births.[25-28] The diagnosis is not a disease, but describes a group of complex neurological disorders acquired early in life (i.e. prenatal, perinatal, or postnatal) with multiple causes and patterns of permanent impairments in the brain.[29] Etiologies can include intrauterine infections, multiple gestation, placental pathology, intrauterine growth restriction, preterm birth, low birth weight, neonatal stroke, neonatal encephalopathy, or CNS congenital malformations.[30-32,29]

2.3.1 **Definition**

The disorder was first described in the mid-18th century by Dr. William John Little.[33] Since then, several definitions and classifications of CP have been proposed.[34] The most recent definition was proposed in 2006 by the International Executive Committee for the Definition of Cerebral Palsy:

> *"Cerebral palsy describes a group of permanent disorders of the development of movement and posture, causing activity limitation, that are attributed to non-progressive disturbances that occurred in the developing fetal or infant brain. The motor disorders of cerebral palsy are often accompanied by disturbances of sensation, perception, cognition, communication, and behaviour, by epilepsy, and by secondary musculoskeletal problems."*[35]

This definition highlights the heterogeneity of the disorder, but emphasizes that impaired movement and posture due to a disturbance in the brain is the invariable clinical manifestation. Furthermore, the functional consequences of the movement impairments are highlighted, in addition to the notion that CP is a life-long condition affecting development. This implies that the management of CP should focus on promoting functional abilities that are developmentally appropriate in a lifespan perspective. Thus, when it comes to hand function in children with CP, knowledge of the natural development is important to be able to distinguish between changes in the child's hand function because of increasing age and changes arising in response to treatment.[36]

2.3.2 **Classification of CP subtypes**

Since CP is a heterogeneous condition covering a wide range of clinical manifestations and degrees of activity limitations, there is a need to classify individuals with CP into more homogeneous subgroups for clinical and research purposes. In general, CP is broadly categorized into three subtypes based on the nature of the predominant type of motor disorder; spastic, dyskinetic or ataxic.[35,37] Spasticity can briefly be defined as increased muscle tone and pathological reflexes, while dyskinesia usually is characterized by fluctuating tone, and involuntary, uncontrolled and recurring movements. Furthermore, ataxia is commonly characterized by low tone, and by imprecise and uncoordinated movements.[37] Spastic CP is the most common subtype affecting more than 80% of individuals with CP.[26,38,39,28] This subtype is subdivided into bilateral when both sides of the body are involved and unilateral when one side is involved.[37] The dyskinetic and ataxic CP subtypes are less common, occurring in approximately 7% and 4% of individuals with CP, respectively.[28,26,40] These latter subtypes usually have a bilateral motor involvement, but their anatomical distribution of motor disorders is rarely denoted. In this thesis, the anatomical distribution is used also for children with dyskinetic and ataxic CP. Consequently, the term bilateral CP includes all children with involvement of both sides of the body regardless of predominant type of motor disorder (i.e. spastic, dyskinetic or ataxic).

2.3.3 **Functional classifications**

Functional classifications are also required to provide common terminology that can ensure consistency when describing functioning and limitations, predicting potential future status, comparing different cohorts (e.g. in population based studies or intervention studies), and monitoring the individual child with CP at different time points.[35] In this way, classifications can enhance communication between clinicians, researchers and families on understanding the child's functional abilities, setting goals and making management decisions.[41] Moreover, a classification system with predictive validity can help parents to anticipate their child's future functioning, and can assist in the identification of best practice for children who perform at different levels. This is crucial for prognostic counselling and for appropriate planning and evaluation of interventions.[42,35,43]

A number of classification systems describing functional limitations seen in children with CP have been developed, including the Gross Motor Classification System (GMFCS)[44] classifying gross motor performance (i.e. sitting and walking), the Manual Ability Classification System (MACS)[45] classifying manual performance (i.e. how children handle objects in daily life), the Bimanual Fine Motor Function (BFMF)[46] classifying fine motor capacity (i.e. ability to grasp, hold, and manipulate), and more recently the Communication Function Classification System (CFCS)[47] classifying capacity to communicate within real-life situations, and the Eating and Drinking Ability Classification System (EDACS)[48] classifying eating and drinking performance.

The usefulness of a classification depends on whether or not the different levels are meaningful and understandable, and have clear descriptions.[45] Thus, appraisal of measurement properties, such as validity and reliability, is required to ensure that the classification provides a similar and consistent understanding between and within users. The

MACS has been found to be a reliable, stable, and valid classification of manual performance and is extensively used worldwide both in clinical practice and in research.[45,49-52] The BFMF is currently the classification system used in the common data base of CP registers established by the Surveillance of Cerebral Palsy in Europe (SCPE),[53] and has been used in a number of epidemiological studies.[26,46,54,38,55] The intention of the BFMF is to classify fine motor capacity, not performance (Beckung, personal communication), and although this was not clearly stated in the original description of the BFMF, this was communicated to the collaborating CP registers in Europe (SCPE).[46] Inter-rater reliability of the BFMF was recently found to be high in a small study including 20 participants,[56] but the classification system had not yet been validated. In addition, it has been indicated that the administration guidelines and the descriptions of the BFMF need to be clarified.[56] Thus, further appraisal of measurement properties, such as validity and reliability, was required to ensure that a similar and consistent understanding of the BFMF was evident between and within users of the classification.

2.4 Assessing hand function in children with cerebral palsy

As described above, functional classifications, such as the MACS and the BFMF, are used in various studies to increase our understanding of distribution and characteristics of functional limitations in populations of children with CP.[2] In addition, the systematic use of assessments, such as classification systems and standardized outcome measures, is a basic requirement for collecting evidence about history of development and treatment effects in the upper extremities of children with CP.[20,57] The use of classification systems and outcome measures have differing purposes. As previously stated, classifications are meant to categorize and discriminate, while outcome measures can assess and describe more precisely a child's hand function.[58] Thus, classifications should not be used to evaluate the impact of intervention, since the classifications do not have appropriate sensitivity for this purpose.[57] By contrast, outcome measures are used to evaluate changes that occur over time, usually associated with either the impact of a health condition (i.e. CP) on the child's development, or the effectiveness of treatment.[59]

2.4.1 **Standardized outcome measures**

Evaluating change as a consequence of treatment is one of the primary reasons for assessing hand function in children with CP. However, when knowledge regarding the natural history of development is limited, it is difficult to establish if changes that follow treatment are the result of the treatment or are just "natural history".[60] Longitudinal studies and outcomes research provide more realistic information about the expected developmental trajectories, and enable therapists to counsel families about the expected outcomes for their child. For the child and their families, the use of individualized outcome measures can provide objective information about strengths and limitations regarding the child's functional abilities and can also demonstrate changes in ability over time. This may motivate the child and the family to increased efforts to gain improvement in the areas measured on the assessment.[61] To be useful, however, the selected outcome measures need to be directly related to the treatment targets, be feasible and practical to administer, be reliable, valid and responsive to change for the population of interest, and most importantly be relevant to the children being assessed and their families.[61]

It is therefore essential that we know what we want to accomplish before beginning to measure, and that we use outcome measures that will enable us to measure that particular task. In general, measurement of functioning has the following three intentions: (1) to discriminate between children for diagnostic or prognostic purposes, (2) to screen for children who are suspected to have developmental problems, or (3) to evaluate changes in functioning over time. [62,63] Discriminative tests are most often norm-referenced and provide information about a child's development compared to age-related typical development. Screening measures are intended to identify children who will require additional in-depth testing, and do not evaluate functional abilities in detail. Evaluative measures are often criterion-referenced, meaning that the child's performance is assessed and scored according to preset criteria, rather than comparing the performance with age-related norms. When choosing an outcome measure, it is therefore important to use a measure that has been created and validated specifically for our intended purpose.[62,57] Thus, if we want to be able to evaluate change over time, then evaluative measures that are criterion-referenced will most often serve this purpose.

In addition to considering the intended purpose of outcome measures, we need to consider whether or not the measure in question meets the needs of the child. For children with CP, the large variety of motor impairments and additional comorbidities represent a challenge when assessing hand function, causing extremely heterogeneous degrees of functional limitations.[60] In addition, functional use of the hands may be influenced by personal factors, such as motivation, attention and concentration, as well as the environmental context (e.g. physical enablers and barriers) in which the manual task is performed.[24] Thus, it can be a challenge to select the most useful and appropriate outcome measure to assess the child's strengths and limitations. This may be difficult even when one is aware of the available outcome measures.[57] A potential solution is to perform a systematic review of outcome measures and thereby facilitating the selection of instruments.[63] Recent reviews have confirmed that valid and reliable outcome measures exist to evaluate hand function in children with unilateral cerebral palsy.[64-68] However, for children with bilateral CP, no such review exists. Thus, it is of upmost importance to identify and appraise outcome measures that evaluate hand function also in these children.

The selection of appropriate outcome measures can be facilitated by systematic appraisal of the content and measurement properties of the instruments in question.[69] An outcome measure used for evaluation should address the domain of concern, be valid (i.e. measure what it is supposed to measure) and reliable (i.e. be consistent) in the population of interest, be responsive to change (i.e. able to demonstrate change), and have good clinical utility (i.e. appropriate, accessible, practical and acceptable).[70,71] Various tools have been developed to aid in the appraisal of the content and measurement properties of outcome measures. For example, the ICF-CY framework can be used to explore and compare the content of various instruments by linking the meaningful concepts to relevant ICF-CY categories.[4,72] Furthermore, the CanChild Outcome Measures Rating Form can be used to extract general characteristics of the outcome measures, such as focus, scale construction, standardization and clinical utility.[73] In addition, the methodological quality of studies investigating measurement properties of selected outcome measures can be evaluated according to criteria described by the COnsensus-based Standards for the selection of health Measurement INstruments (COSMIN).[69] Moreover, the results of the measurement properties for each outcome measure can be rated according to the quality criteria proposed by Terwee and colleagues.[74]

2.4.2 **Outcome measures assessing bimanual performance**

Improved functional performance in daily life is the ultimate goal of most interventions. Still, a large proportion of the available hand function measures focus on capacity to use the hands, rather than on actual use of the hands when handling objects in daily life.[57] In addition, most of the hand function measures evaluate one hand at a time, such as the Melbourne Assessment 2 (MA2) and the Quality of Upper Extremities Skills Test (QUEST), while most activities performed in daily life require the use of both hands together.[57,75,76]

For children with unilateral CP the Assisting Hand Assessment (AHA) has been developed to measure how effectively the affected hand is used while performing bimanual tasks.[7] The AHA has proven to produce valid and reliable outcome measures.[11,8,10,9] In addition, the test has been found to be responsive to change, allowing monitoring of upper extremity development and evaluation of interventions in children with unilateral CP.[13,11] The test is criterion-based, and the scale construct has been refined and evaluated by the means of Rasch measurement model analysis.[77] The use of the Rasch model is the current standard for the development of unidimensional scales, contributing with metric quality outcomes in healthcare.[78] By using the Rasch model analysis, ordinal raw scores can be transformed into interval scale measures (log-odds probability units) and the internal validity of the scale and aspects of reliability can be investigated.[77] Interval scale measures are recommended for reporting outcomes that are used to evaluate change.[78] In addition, the Rasch analysis creates item difficulty hierarchies that can be used to generate an ability profile for the individual child, identifying the items the child performs well, as well as items that are not yetaccomplished, but are close to the child's next ability level. The latter items may represent "just the right challenge" for the individual child and can indicate possible targets for intervention.[10] In this way the Rasch-derived AHA measure may contribute to improve goal-setting in children with unilateral CP.

For children with bilateral CP there is no equivalent performance-based outcome measure with a bimanual perspective.[57] Knowing that children with bilateral CP have varying degrees of impairments in both hands that may severely influence their ability to successfully perform everyday activities, it is important to develop bimanual performance measures that are validated for children with bilateral CP.[6] Thus, it would be of great interest to explore if the AHA could be modified for use with children with bilateral CP of spastic, dyskinetic and ataxic type.

2.5 Hand function in children with cerebral palsy

Hand function in children with CP has been described in several population-based studies using the MACS classification system.[5,49,79-81] These studies have indicated that more than 60% of the children can be considered to be independent in the performance of most age-related daily activities (MACS levels I-II), while nearly 40% need different levels of assistance (MACS levels III-V). However, being independent does not imply that the child with CP has no problems with his or her hand function. Children at MACS level I (39-42%) handle most objects easily and successfully, but may experience limitations in new situations or when advanced skills are required. Children at MACS level II (19-22%) handle most objects, but with somewhat reduced ability and/or speed. These children may avoid certain activities or the activities are achieved with slowness and reduced quality of performance.[45]

Children classified at MACS level III (9-14%) commonly require help in preparing and/or modifying activities to be able to perform activities independently. These children can handle objects, but with difficulty. Thus, their degree of independence is related to the degree of supportiveness of the environmental context.[45] Children classified at MACS level IV (8-12%) and V (12-24%) have very limited or no ability to handle objects. These children are dependent on continuous help and support and can at best participate meaningfully in parts of an activity, or with simple movements in special situations (i.e. pushing a button).[45]

The BFMF classification of fine motor function has also been used in population-based studies to describe hand function.[26,46,38,80] According to these studies more than 70% of the children had at least one hand that could manipulate without restrictions or had limitations only in more advanced fine motor skills (BFMF I-III), while nearly 30% could only grasp or worse with both hands (BFMF IV-V). The children who had one hand with good or relatively good ability, had increasingly impaired fine motor function in the other hand as indicated by their BFMF levels (BFMF I: 31-45%, BFMF II: 23-33% and BFMF III: 5-12%). Among the children who could only grasp or worse with both hands, 7-12% of the children could grasp some objects (BFMF IV), while 14-18% could only hold or worse (BFMF V).

In addition to population-based descriptions of hand function in the heterogeneous group of children with CP, more detailed information is required to enable monitoring of upper extremity development and evaluation of interventions. In this thesis, what is known regarding functional use of the hands in children with CP will be described separately for children with bilateral CP of spastic, dyskinetic and ataxic types (included in studies I-IV) and children with unilateral CP (included in studies I and II).

2.5.1 **Bilateral CP of spastic, dyskinetic and ataxic type**

Bilateral spastic CP accounts for approximately 55% of the whole population of children with CP, while dyskinetic CP is evident in about 7% and ataxic CP in about 4%.[28] In children with bilateral CP, the upper and/or the lower extremities on both sides of the body, as well as the posture, are affected to a greater or lesser degree. Children with these subtypes have been found to be distributed between all GMFCS, MACS and BFMF levels, with somewhat fewer children with ataxic CP classified at the higher levels.[5,82,49,38] Children classified in GMFCS levels III-V usually require increasing amounts of postural control in sitting to facilitate reaching.[83] Nearly 70% of the children can handle objects, although with varying difficulty (MACS levels I-III), while more than 30% of the children have been found to have very limited or no ability to handle objects (MACS levels IV-V).[5] Whereas some of these children have two relatively well-functioning hands, others have low functional abilities in both hands. Moreover, some children may have a relatively asymmetric hand function where one hand is clearly more affected than the other.[5] Common features influencing the functional use of the hands differ somewhat according to the type of motor impairment. Bilateral spastic CP is generally characterized by velocity-dependent increase in muscle tone, muscle weakness, slow and stiff movements, decreased range of motion, impaired motor control and lack of coordination,[84,85,5,53,86-88] while dyskinetic CP is typically characterized by slowness, fluctuating tone, and involuntary, uncontrolled, recurring and occasionally stereotyped movements that increase during activity.[89,90,53,91] Moreover, ataxic CP is commonly characterized by low tone, and by imprecise and uncoordinated movements.[53,37,92] In several children with bilateral CP, a co-existence of motor impairment types may be present.[53,93,94,92]

Additional impairments may further complicate the functional use of the hands. Children with bilateral spastic and dyskinetic CP have been found to experience tactile deficits.[95-97] Furthermore, population-based studies have shown that comorbidities such as visual impairments (bilateral spastic: 28-44%, dyskinetic: 21-45%, ataxic: 8-17%),[46,98,38,99] learning or intellectual disability (bilateral spastic: 33-58%, dyskinetic: 52-77%, ataxic: 50-80%),[98,46,38,80,100] , behavioral problems (bilateral spastic: 48%, dyskinetic: 11%, ataxic: 13%),[101] or epilepsy (bilateral spastic: 30-62%, dyskinetic: 42-70%, ataxic: 7-40%)[38,46,98,26] are common. Children with more severe manual impairments are more likely to also have severe visual and intellectual impairment.[101]

Impaired hand function in children with bilateral CP has been found to influence their ability to successfully perform daily activities.[8,9] Most activities performed in daily life require the use of both hands. Still, little is known about how children with bilateral CP use their hands together when handling objects. One reason for this may be the traditional focus on gross motor function and facilitating of normal movement patterns in these children, with the assumption that gross motor abilities is a prerequisite for fine motor skills.[36,102] However, intervention studies targeting the upper extremities using this approach have shown little improvement.[36,103,104]

Among children with bilateral CP, a large variation in bimanual hand use and the pattern of development of hand function can be expected because of the large variation in motor impairments and additional impairments.[105] In a literature search, no studies investigating the longitudinal development of hand function in children with dyskinetic or ataxic CP were identified, while three studies were identified for children with bilateral spastic CP.[105-107] The first study suggests that children with bilateral spastic CP have limited potential for improvement of movement patterns measured with the QUEST,[76] while fine motor skills measured with the Peabody Developmental Motor Scales (PDMS)[108] showed some improvement from 16 months, but with a decline already from 3 years of age. The second study reported that the speed and movement efficiency in grasping with the dominant hand in children with bilateral spastic CP improved during a 13-year period measured with the Jebsen-Taylor test of hand function[109] and a precision grasping task.[106] Similar results were reported in the third study, showing that finger grip strength and finger movement velocity improved in the dominant hand in pre-school aged children with bilateral spastic CP within a one-year period.[107] Thus, the results from the three studies suggest that functional skills in the dominant hand probably improve by age, while movement patterns in the upper limbs may decline. However, there are still knowledge gaps concerning the natural history of development of bimanual hand use in children with bilateral CP, as well as regarding effects of treatment targeting the upper extremities in these children.[103]

2.5.2 Unilateral spastic CP

Unilateral spastic CP has been found to account for more than 30% of the whole population of children with CP,[28] and is characterized by sensorimotor impairments of one side of the body. Thus, these children in general have one well-functioning hand and one affected hand.[110] Children with this subtype are commonly able to walk with or without aids (GMFCS I-III),[26,49,38] and they usually present with manual ability and fine motor function corresponding to MACS and BFMF levels I-III.[5,38,26] The functional ability in the affected hand varies from showing only some clumsiness in tasks requiring high precision and manipulative skills, to having no ability to grasp or hold.[111,11] Common impairments influencing the functional use of the affected hand are increased muscle tone, muscle weakness, decreased range of motion, slowness, impaired selective motor control, and coordination difficulties, which occur to a varying extent in the children independently of age.[112,113,5,110,114-119] In addition to the motor problems, several children with unilateral CP experience tactile deficits (75-90%)[120,113] or visual impairments (8-11%).[99,46,38] Furthermore, intellectual or learning disability (5-19%),[80,38,46,121,100] behavioural problems (24%)[101,122] or the presence of epilepsy (19-26%)[123,26,38] and pain (51%) are relatively common in children with unilateral CP. All of these additional impairments may also influence the ability to handle objects.[7]

Usually children with unilateral CP can easily manage activities requiring the use of only one hand. However, several activities performed every day involve the use of both hands, and difficulties in reaching, grasping, manipulating and releasing objects with the affected hand will therefore affect the performance of these activities.[115] Recent studies using tests such as the AHA and the Hand Assessment for Infants (HAI) have identified development of hand function and aspects of how children with unilateral CP use the affected hand spontaneously in the performance of bimanual tasks.[7,11,111,10,13,124] These studies indicate that the amount and quality of how the two hands are used develops differently already at an early age. Asymmetric hand use can be seen already at the age of 3–5months.[124] As these children develop they tend to acquire more and better skills with the well-functioning hand compared with the affected hand.[125,13] Although the functional skills of the most affected hand never catch up with the well-functioning hand, a steady increase in functional abilities of the affected hand during preschool years has been identified. This increase has been found to be more rapid and favourable in children who are able to grasp objects already at 18 months of age, compared with children who have difficulties grasping objects at the same age.[13] Curves describing this predicted development of hand function have been published, indicating that most of the development takes place before 3-4 years of age in the high-ability group, and before 7 years of age in the low ability group.[13] However, the functional use of the hands can be improved also after preschool age,[106] even in children with more severe impairments, by learning appropriate strategies for using the affected hand in bimanual tasks.[20,11]

2.6 Summary of the background and the identified "knowledge gaps"

Among children within the heterogeneous diagnosis CP, a large variation in functional use of the hands can be observed. To be able to describe common characteristics of hand function in populations of children with CP, it is necessary to have classification systems that can be used to describe hand function in a valid and consistent matter. The BFMF is currently the classification system used in the common data base of CP registers established by the SCPE, and has been used in a number of epidemiological studies. Inter-rater reliability of the BFMF was recently found to be high in a small study including 20 participants,[56] but the classification system has not yet been validated. To ensure consistent understanding and interpretation of the BFMF, there is a need to validate measurement properties of this classification system, as well as investigating both intra- and inter-rater reliability with larger samples.

In addition to classifications giving broad descriptions of hand function, there is a need for more detailed outcome measures that are able to measure and describe the development of hand function children with CP and to evaluate interventions.[57] For this purpose, valid and reliable outcome measures that are responsive to change are crucial.[70] The selection of appropriate outcome measures can be facilitated by systematic appraisal of the content and measurement properties of outcome measures.[69] Recent reviews have confirmed that valid and reliable outcome measures exist to evaluate hand function in children with unilateral cerebral palsy.[64-68] However, for children with bilateral CP, no such review exists. Thus, it is important to identify and appraise outcome measures evaluating hand function also in these children.

Furthermore, little is known regarding how children with bilateral CP use their hands when handling objects in bimanual activities. One reason for this may be the lack of available outcome measures with a bimanual perspective. Knowing that children with bilateral CP have varying degrees of impairments in both hands that may severely influence their ability to successfully perform everyday activities, it is essential to develop and validate outcome measures evaluating bimanual performance.[6] Thus, it would be of great interest to explore if the AHA could be modified for use in children with bilateral CP, so that bimanual performance could be evaluated and described also in these children.

3. AIM

The overall aim of this thesis was to develop and validate assessments (i.e. classifications and outcome measures) of hand function in children with bilateral CP, to enable measurement of change and generalization of results. This has been explored in four papers with the following specific objectives:

Paper I: To examine aspects of construct and content validity of the Bimanual Fine Motor Function (BFMF) classification in children with CP.

Paper II: To present a revised edition of the BFMF, emphasizing it as a classification of fine motor capacity and including a clear description separating the five levels. Furthermore, the aim was to explore intra- and inter-rater reliability of the new edition in children with CP.

Paper III: To review outcome measures used to evaluate hand function, with emphasis on manual capacity and performance, in children with bilateral CP, to describe the content and measurement properties of such measures, and to investigate the quality of the studies that have examined these properties.

Paper IV: To develop and validate a new test for children with bilateral CP, the Both Hands Assessment (BoHA), which could measure bimanual performance, as well as quantify a possible side difference between hands.

4. METHODS

4.1 Study outline

An overview of the study outline of the four papers included in this thesis is shown in Table I. The first two papers investigated measurement properties of the BFMF classification in children with CP. Paper I explored construct and content validity of the BFMF, and paper II investigated reliability of a revised edition of the BFMF. In this edition, the original wording of the BFMF classification levels was retained. In addition explanatory figures with more precise descriptions of the levels were added and it was emphasized that fine motor capacity should be classified. Paper III was a systematic literature review of outcome measures used to assess hand function in children with bilateral CP. Paper IV was a test development study with a cross-sectional design. In this study, the content of the new test BoHA was developed through adaptation of the AHA and the Rasch measurement model was used in the development and validation of the test.

Table I. Overview of the four studies described in papers I-IV of this thesis

Characteristics of the studies	Paper I	Paper II	Paper III	Paper IV
Study design	Register-based Cross-sectional Content comparison using ICF-CY as frame of reference	Refinement of instrument Cross-sectional	Systematic review	Test development Cross-sectional
Statistical analysis	Spearman's rho	ICC Absolute agreement Cohen's weighted kappa		Internal scale validity Item and person reliability Person separation ratio Spearman's rho Pearson correlation
Qualitative methods	Content analysis Linking to ICF-CY	Content analysis	Descriptive Quality assessment of studies and outcome measures	Content analysis

ICF-CY=International Classification of Functioning, Disability and Health Children and Youth version, Spearman's rho=Spearman's rank correlation, ICC=Intraclass Correlation

4.2 Study population

In this section, details regarding populations included in papers I, II and IV are presented initially (Table II), before information regarding outcome measures and papers included in the systematic review are presented (Figure 3).

In the construct validity part of paper I, a total of 539 children (304 boys and 235 girls) with CP were included from the CP register of Norway (n=384) and western Sweden (n=155). Only children with a valid classification of hand function by both the BFMF and the MACS were included. This comprised 55.1% of the children with CP born in Norway in 1999-2003 and 83% of the children with CP born in 1999–2002 and living in western Sweden at census date. Median age at recording in the Norwegian register was 6 years (range 4 to 12 years), while it was 5 years (range 4 to 8 years) in the Swedish register. The included children represented all CP subtypes and MACS, BFMF and GMFCS classification levels (Table II).

Table II. Characteristics of the study populations

	Paper I	Paper II	Paper IV
Number of participants	539	79	171
Norway	384	36	80
Sweden	155	43	91
Age, range years	4-12	3-17	1.5-12
Gender			
Boys	304	45	96
Girls	235	34	75
CP subtype			
Unilateral spastic	229	19	
Bilateral			171
Spastic	218	42	
Dyskinetic	57	15	
Ataxic	29	3	
Unspecified	6		
MACS			
I	179	24	53
II	180	27	56
III	60	16	55
IV	36	5	
V	84	7	
Unclassified			7
BFMF			
I	217		
II	160		
III	53		
IV	33		
V	76		
GMFCS			
I	267	28	
II	88	28	
III	37	8	
IV	53	7	
V	94	8	

CP=cerebral palsy, MACS=Manual Ability Classification System, BFMF=Bimanual Fine Motor Function, GMFCS=Gross Motor Function Classification System

For paper II, a convenience sample of 79 children (45 boys and 34 girls) with CP was recruited at two study sites; the Regional Rehabilitation Centre, Queen Silvia Children's

Hospital, Göteborg, Sweden (n=43) and St Olav Hospital in Trondheim, Norway (n=36). The children who were recruited in Norway also participated in paper IV. The mean age of all the children was 8 years and 7 months (range 3 to 17 years) and the children represented all CP subtypes and all MACS and GMFCS levels, see Table II.

For paper IV, the total study population comprised a convenient sample of 171 children (96 boys and 75 girls) with bilateral CP of spastic, dyskinetic and ataxic type. The children were recruited through pediatric habilitation units in Sweden (n=91) and Norway (n=80) and their mean age was 6 years and 6 months (range 1 year and 6 months to 12 years), see Table II. Only children with hand function corresponding to MACS levels I-III were included, since children at these levels can be expected to be able to handle the objects used in the BoHA test situation, although with varying degrees of difficulty. Seven of the children were not classified according to MACS in their medical records. Children classified to MACS levels IV-V were excluded, since they by definition have very limited, or no, ability to handle objects.[45]

In the systematic review described in paper III a total of 212 full-text papers were assessed for eligibility, of which 16 of the papers and five hand function measures were included for quality assessment, see Figure 2. The included outcome measures are described in Table VII in the Results section 5.2.1.

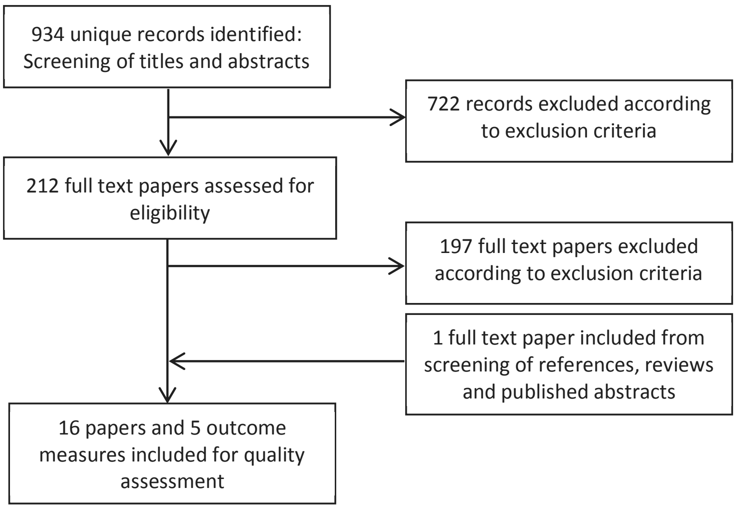


Figure 3: Processes performed to identify hand function measures and studies of measurement properties.

Outcome measures were included if (1) the primary intention was to evaluate hand function according to how children handle objects, i.e. within the activity component, or a combination of the activity and body function components, of the ICF;[23] (2) measurement properties of the outcome measure had been evaluated in children with bilateral CP in the age range 0-18 years; and (3) the measure was designated for use in a clinical setting, without the need for laboratory equipment. Papers investigating measurement properties were excluded if they were: (1) not published in English; (2) evaluated measurement properties of classification systems or of measures primarily assessing participation, body function, body structures, school-functioning, self-care or health-related quality of life; (3) children with bilateral CP comprised less than 30% of the total population; or (4) they were not a full-text original article.

4.3 Data collection

Data for papers I, II and IV were collected using classifications and outcome measures assessing hand function (Table III), while systematic search strategies and quality assessments were used to collect data for paper III.

Table III. Overview of included classifications and outcome measures.

	Paper I	Paper II	Paper IV
MACS	X	X	X
BFMF	X		
BFMF 2		X	
GMFCS	X	X	
CP-subtype	X	X	X
BoHA			X

MACS=Manual Ability Classification System, BFMF=Bimanual Fine Motor Function, BFMF 2=Bimanual Fine Motor Classification second edition, GMFCS=Gross Motor Function Classification System, CP=cerebral palsy, BoHA=Both Hands Assessment

For paper I, the CP registers of Norway and western Sweden provided clinical data describing the children's BFMF and MACS classification levels, as well as CP subtype and GMFCS level. Clinicians working at the public habilitation centers caring for the children collected the data. None of these professionals were aware of the study question.

For paper II, four raters independently classified the children's fine motor function from video recordings applying the BFMF 2 at two different time-points at least two weeks apart. Three of the raters had developed the BFMF 2. In addition, a fourth rater was included who was not familiar with the BFMF. In addition to the BFMF classifications, information on CP subtype, and MACS and GMFCS classification levels was obtained from medical records.

For paper IV, bimanual performance in children with bilateral CP, in the age range between 18 months and 12 years, was assessed using the new test BoHA. In addition, information on MACS classification levels was obtained from medical records for each child. The BoHA test sessions were administered and video-recorded by examiners who were certified AHA-raters. The scorings were done by the BoHA test developers.

For paper III, searches for outcome measures were carried out in Embase, MEDLINE, PubMed and CINAHL until 10 June 2015. The first search aimed to identify hand function measures used in children with bilateral CP (details of the database search are presented in paper I, Appendix 1). Subsequently, the names of the outcome measures identified through the first search were used in a complementary search, which aimed to identify additional studies of the measurement properties of the hand function measures.

4.3.1 **Classifications and outcome measures**

4.3.1.1 Bimanual Fine Motor Function (BFMF)

The BFMF classifies fine motor function in children with CP.[46] Five ordinal grading levels are used with lower levels indicating higher ability (Table IV). The classification level is determined by assessing the child's capacity to grasp, manipulate, and hold objects for each hand separately. The classification levels II–IV can be further subdivided into (a) and (b) according to variations between the two hands in capacity to grasp, hold and manipulate.[46] Inter-rater reliability of the BFMF was recently found to be high,[56] but the classification system had not yet been validated.

Table IV: The five classification levels of the Bimanual Fine Motor Function (BFMF). Text in italics shows added text for the modified BFMF version used in the cerebral palsy registers in Norway and western Sweden.

BFMF
Level I One hand: manipulates without restrictions. The other hand: manipulates without restrictions or limitations in more advanced fine motor skills.
Level II (a) One hand: manipulates without restrictions. The other hand: only ability to grasp or hold. (b) Both hands: limitations in more advanced fine motor skills.
Level III (a) One hand: manipulates without restrictions. The other hand no functional ability. (b) One hand: limitations in more advanced fine motor skills. The other hand: only ability to grasp or worse. *The child needs help with tasks.*
Level IV (a) Both hands: only ability to grasp. (b) One hand: only ability to grasp. The other hand: only ability to hold or worse. *The child needs support and/or adapted equipment.*
Level V Both hands: only ability to hold or worse. *The child requires total assistance, even with adaptations.*

A modified version of the BFMF (see Table IV) was used in the CP registers of Norway and western Sweden whereby the child's need for help or adaptations was added for levels III–V. Furthermore, the BFMF was classified without subdividing the levels into (a) or (b). This modified BFMF version was used in paper I of this thesis, similar to the version used in the European SPARCLE study.[54]

For paper II, a revised edition of the BFMF (BFMF 2) was developed and used (see paper II, Appendix S1). In this edition it was emphasized that fine motor capacity should be classified. In addition, explanatory figures with more precise descriptions of the classification levels were added and used together with the original wording of the BFMF classification levels. To enable investigation of intra- and inter-rater reliability of the BFMF 2, the assessments of the children's fine motor capacity were video-recorded. The children were asked to do tasks with each hand separately involving grasping and holding of three objects of different size (pencil, cereal/raisin/non-stop and cube), as well as in-hand-manipulation of the cube.

4.3.1.2 Manual Ability Classification System (MACS)

The MACS classifies how children with CP use their hands to handle objects in daily activities, classifying usual performance. The classification system has been found to be valid and reliable for children with CP, 4-18 years of age,[45,49-52] and describes five distinct ordinal levels of manual performance, with lower levels indicating higher ability (see Table V).

Table V: The five classification levels of the Manual Ability Classification System (MACS). The complete MACS brochure can be downloaded from http://www.macs.nu/

MACS
Level I Handles objects easily and successfully. At most limitations in the ease of performing manual tasks requiring speed and accuracy. However, any limitations in manual abilities do not restrict independence in daily activities.
Level II Handles most objects, but with somewhat reduced quality and/or speed of achievement. Certain activities may be avoided or achieved with some difficulty; alternative ways of performing might be used, but manual abilities do not usually restrict independence in daily activities.
Level III Handles objects with difficulty; needs help to prepare and/or modify activities. The performance is slow and achieved with limited success regarding quality and quantity. Activities are performed independently if they have been set up or adapted.
Level IV Handles a limited selection of easily managed objects in adapted situations. Performs part of activities with effort and limited success. Requires continuous support and assistance and/or adapted equipment for even partial achievement of the activity.
Level V Does not handle objects and has severely limited ability to perform even simple actions. Requires total assistance.

The MACS levels are determined based on the children's self-initiated ability to handle age-appropriate objects in daily activities and their need for assistance or adaptation to perform manual activities in everyday life at home, school, and community settings. Thus, information from a parent or someone who knows the child well is important in the selection of the most appropriate classification level. A MACS identification chart can be used to guide the decision process (available at http://www.macs.nu/level-identification-chart.php).

4.3.1.3 Both Hands Assessment (BoHA)

The BoHA measures bimanual performance in children with bilateral CP in the age range 18 months to 12 years, with a hand function corresponding to MACS levels I-III. The AHA test-kit is used in the BoHA test situation to elicit spontaneous collaborative use of the hands in playful tasks, so that bimanual performance can be scored. For children aged 18 months to 5 years, the Small Kids AHA test-kit was applied using explorative play, while for 6 to 12-year-old children the School Kids AHA test-kit was used with board games as the age-appropriate test session.[10] Furthermore, the AHA set-up for administration and video-recording was followed except for one adjustment: In the BoHA test situation, the toys were placed on both sides of the child equally often, as opposed to the AHA test situation in which most of the objects are to be placed at the child's affected hand side.[7]

Table VI. Overview of the items included in the Both Hands Assessment (BoHA) scale.

BoHA
Unimanual items scored separately for the dominant (D) and the non-dominant (ND) hand
Initiation items
1. Initiates use
Movements items
2. Speed of movements
3. Quality of arm movements
4. Quality of finger movements
5. Reaches
Grasp-release items
6. Grasps
7. Stabilizes objects
8. Varies type of grasp
9. Releases
Fine motor adjustment items
10. Grip force regulation
11. Manipulates
Bimanual items scored with one common score for both hands together
Coordination items
12. Readjusts grasp
13. Coordinates
14. Orients objects
Pace items
15. Proceeds
16. Flow in bimanual task performance

Based on the observation of bimanual play, the BoHA is scored on 16 items (11 unimanual and 5 bimanual), on a 4-point rating scale (see Table VI and VII). The eleven unimanual items are scored separately for the dominant (D) and the non-dominant (ND) hand. The unimanual sum scores are used to determine a possible difference between the hands, reported as percentage difference; the higher the number, the greater the difference between the sides. For each of the five bimanual items, one common score is given for both hands. This score is added to the unimanual scores, resulting in a total of 27 data points summed up as the "Both hands sum-score", a raw score reflecting overall bimanual performance.

Table VII. Examples of 4-point rating scales for one unimanual item that is scored for the dominant (D) and the non-dominant (ND) hand, and for one bimanual item scored with one common score for both (B) hands.

Item examples	Scoring criteria describing the 4-point rating scale			
Unimanual item			D	ND
3 Quality of arm movements	Easily performs accurate and balanced movements of the assessed arm when approaching and handling objects.	4		
	Arm movements are only somewhat imprecise or exaggerated, not obviously affecting performance.	3		
	Arm movements are often involuntary or exaggerated or inaccurate; or arm movements are performed with effort; or the arm is often held close to the body, i.e. there is little variation of arm movements.	2		
	Most often arm movements are involuntary and/or highly excessive, implying task break down; or does not change the position of the arm; or on single occasion changes the position of the arm.	1		
Bimanual item			B	
12 Readjusts grasp	Often, easily and automatically re-grasps objects, often readjusts grasp.	4		
	Questionable ability or slight delay in changing/readjusting grasp, but re-grasps relatively often; or the child relatively often moves objects between his/her hands, but with some delay.	3		
	Seldom re-grasps an object or has difficulties in changing/readjusting grasp, i.e. most often holds the objects with the same grasp throughout the task.	2		
	Does not change/readjust grasp even though it is ineffective, which results in breakdown of some task, or does most often not use grasp.	1		

In addition to the 16 BoHA items, the child's dominant hand is noted as well as the child's level of postural control and need for support in sitting.

4.3.1.4 Other variables

Information on gross motor function classification levels (GMFCS), as well as information regarding CP subtypes, was obtained from hospital medical records for each child.

4.4 Ethical considerations

All studies in the thesis were approved by the Regional Ethical Committee for Medical Research in Mid-Norway. In addition, the first and second studies were approved by the Ethics Committee at the Medical Faculty at Gothenburg University, Sweden, and the fourth study was approved by the Ethics Research Committee of Karolinska Hospital in Stockholm, Sweden. Furthermore, the first study was also approved by the Norwegian Data Inspectorate.

Informed consent was obtained from the parents of the children.

4.5 Data analysis

A variety of different statistical and qualitative methods were used to analyse the data included in the four papers (see Table I). The methods used were chosen based on the aim of the papers. For paper I, statistical methods were used to investigate construct validity, while content validity was appraised by qualitative methods. For paper II, agreement was investigated by statistical methods, while qualitative analyses were used to appraise measurement properties of included outcome measures for paper III. For paper IV, qualitative methods were used to develop and validate the content, and statistical methods were used to refine the test and to investigate internal scale validity and aspects of reliability.

4.5.1 **Statistical methods**

Construct validity of the BFMF was assessed by investigating the relation between children's BFMF and MACS ordinal classification levels using the non-parametric Spearman's rank correlation (rho). Furthermore, the non-parametric Marginal Homogeneity test was used to study whether or not there were statistically significant differences in combinations of classification levels between the BFMF and the MACS in the same population.

To investigate reliability of the BFMF, an overall intraclass correlation coefficient (ICC) was calculated using a mixed model with the rating level as dependent variable, the test sessions as categorical covariate, and the children and the raters as crossed random factors. In addition, intra- and inter-rater reliability was analyzed in terms of absolute percentage agreement and agreement by Cohen's quadratic weighted kappa (Cohen's weighted kappa: K_w).

The Rasch measurement model was used to refine the BoHA and to evaluate internal scale validity and aspects of reliability.[77] Using this model detailed analysis of how items work within scales, and whether or not their summed score is valid, can be investigated. The basic theoretical assumption of the Rasch model is that persons with greater ability are more likely to accomplish difficult test items than those with less ability, and that all persons have a greater probability of accomplishing easier test items than ones that are more difficult. Based on transformation of ordinal raw scores into interval scale measures (log-odds probability units), the Rasch model orders item difficulty estimates hierarchically from easiest to hardest, as well as person's ability measures from high to low ability. The difference between observed and expected data is used to determine how well test items fit the underlying construct being measured, and to determine how well each test item contributes to the estimation of the person ability measures.[77]

Unidimensionality of the BoHA scale was investigated by item and person goodness-of-fit statistics and principal components analysis (PCA) of the standardized residuals. In addition, testing for differential item functioning (DIF) was used to investigate whether the scale worked in the same way for groups with different characteristics.[77,126] Furthermore, the item and person reliability coefficients were calculated indicating degree of replication of the item hierarchy and the ability of the scale to reliably rank a person's relative measure location, similar to the Cronbach's alpha.[127] In addition to the Rasch measurement analyses, the association between BoHA measures and MACS levels was calculated by Spearman's rho, and the association between the BoHA measures and age, and gender were calculated by the parametric Pearson correlation (r).

Where relevant, 95% confidence intervals (CI) and significance levels (p-values) were reported.

4.5.2 **Qualitative methods**

Content validity of the BFMF was appraised by the use of literature review to judge the relevance of the classification levels for describing fine motor function in children with CP from very mild to very severe disability. In addition, three of the authors of paper I, trained in the ICF-CY, applied the ICF-CY framework and the linking rules proposed by Cieza and colleagues[4,3] to explore and compare the content and the comprehensiveness of the BFMF and MACS.

Expert consensus, including one of the developers of the original BFMF,[46] was used to develop and validate the content of the revised BFMF 2 edition. The wording and content of the different levels of the original BFMF was discussed and evaluated and video recordings of children performing at different levels of fine motor function were scrutinized. This resulted in the creation of explanatory figures and text that were included in the BFMF 2, together with the original description of the BFMF classification levels.

The data extraction and quality assessment of the identified studies and outcome measures in the systematic review consisted of several steps. First, the general characteristics of the outcome measures, such as focus, scale construction, standardization and clinical utility, were extracted into a review table, adapted from the CanChild Outcome Measures Rating Form.[73] Second, the methodological quality of studies of measurement properties of the included outcome measures was evaluated according to the COnsensus-based Standards for the selection of health Measurement INstruments (COSMIN).[69] Following the COSMIN checklist, the methodological quality per measurement property in each study was rated and described according to a four-point rating scale (i.e. poor, fair, good, and excellent).[128] Third, the results of the measurement properties for each study were rated as positive (+), indeterminate (?), or negative (-), according to quality criteria proposed by Terwee and colleagues.[74] Lastly, the level of overall evidence for each outcome measure was estimated by taking into consideration a) the number of studies, b) the methodological quality of the studies and c) the consistency of their results in a similar manner to that recommended by the Cochrane Back Review Group.[129] The possible levels of evidence are “strong”, “moderate”, “limited”, “conflicting”, or “unknown”.

The content of the BoHA test items were generated from observations of bimanual play in children with bilateral CP and through adaptation of the 20 AHA test items. To decide which items to include in the BoHA scale, the items were sorted as follows: (1) “Suitable - no changes required”; (2) “Suitable after adaptation”; and (3) “Not suitable”. Furthermore, new test items were generated based on object-related hand and arm actions observed in the BoHA video recordings, but that were not covered by the original AHA test items. In addition, the clinical relevance and perceived importance of each item for evaluation of bimanual performance in children with bilateral CP was appraised to decide which items to retain in the BoHA scale.

5. RESULTS

Findings from the four papers will be presented together under the following themes: evidence of validity and reliability of the BFMF (papers I and II), and valid and reliable outcome measures evaluating hand function in children with bilateral CP (papers III and IV).

5.1 Evidence of validity and reliability of the BFMF

5.1.1 Construct validity

The results from paper I suggest that the BFMF is a valid classification of fine motor function in children with CP, as indicated by the high correlation between children's BFMF and MACS levels (rho=0.89, CI: 0.86 to 0.91,p<.001). Altogether 415 (77%) children had corresponding classification levels on the BFMF and MACS. Only four children had more than one level difference between the BFMF and MACS classifications. Both overall, and for the various subtypes of CP, there was a trend towards lower levels (higher ability) for the BFMF than the MACS. This trend was significant for unilateral, bilateral (all *p's* < .001), and dyskinetic (*p* = .020), but not for ataxic CP (*p* = .11).

5.1.2 Content validity

The literature review in paper I confirmed the relevance of the items hold, grasp and manipulate to describe increasingly advanced fine motor abilities in children with CP. Furthermore, similarities and differences between the content of the BFMF and the MACS were identified when the classification systems were linked to the ICF-CY. Both the BFMF and the MACS were linked to the mobility chapter and the second level category fine hand use within the activity and participation component of the ICF-CY. In addition, the MACS was linked to 14 other second-level categories, reflecting the broad perspective of manual performance covered by the MACS (i.e. handling objects in all different contexts of daily life). In contrast, only the meaningful concepts of the BFMF were linked to the more detailed third level categories Picking up, Grasping, and Manipulating, while the MACS was not found to classify specific details regarding fine hand use. Finally, the results from paper I indicated that it could not be determined from the written descriptions of the BFMF whether it was a classification of fine motor capacity or performance, although this had been verbally communicated to the collaborating CP registers in Europe (SCPE).

The development of the revised edition BFMF 2 was based on expert consensus. The content analysis resulted in the BFMF 2 where the original description of the BFMF levels was retained, but explanatory figures with more precise descriptions of the classification levels were added and it was emphasized that fine motor capacity should be classified.

5.1.3 **Reliability**

The findings from paper II demonstrated that fine motor capacity could be reliably classified from short video recordings (< 5 minutes) using the revised BFMF 2 edition. The overall ICC, reflecting the between rater, within individual, agreement was 0.86. The mixed effect model indicated that there was no statistically significant shift between test session one and two (p=0.78), and the variance between raters was not statistically significant (p=0.16). Furthermore, the absolute agreement for intra-rater reliability was found to vary from 0.75 to 0.94%, and from 0.63 to 0.76% for the inter-rater reliability. The Cohen's weighted kappa indicated high intra-rater (κ_w: >0.90) and inter-rater (κ_w: > 0.85) reliability.

5.2 Valid and reliable hand function measures for children with bilateral CP

5.2.1 **Identified outcome measures with evaluated measurement properties**

In the systematic review (paper III) a total of 16 papers and five hand function measures were included for quality assessment. The included hand function measures were one parent-reported questionnaire; the ABILHAND-Kids,[130] and four standardized tests; the Erhardt Developmental Prehension Assessment (EDPA),[131] the Peabody Developmental Motor Scales (PDMS),[108] the Quality of Upper Extremity Skills Test (QUEST),[76] and the Melbourne Assessment of Unilateral Upper Limb Function (MA).[132] For the MA and the PDMS revised versions of the measures have been published, the Melbourne Assessment 2 (MA2)[133] and the Peabody Developmental Motor Scales Second Edition (PDMS-2).[134] These versions were slightly different from the original versions and were therefore treated separately in the quality appraisal of the current level of evidence for the measurement properties.

The content and focus of the hand function measures varied considerably (see Table VIII). The strongest level of evidence was found for the MA2, which measured unimanual capacity, and for the questionnaire ABILHAND-Kids, assessing perceived manual ability (Paper III, Table IV). The criterion-referenced MA2 was based on refinement of the original MA through Rasch modelling and included four discrete subscales (range of motion, accuracy, fluency and dexterity).[75] Thus, the measurement properties of the original MA do not apply to the MA2, and further research is required to determine additional aspects of reliability, as well as responsiveness of the MA2.[135] The ABILHAND-Kids has also been developed and validated through Rasch modelling, and the results indicated strong evidence for internal scale validity and person reliability.[130] However, further investigation of other aspects of reliability and responsiveness to change is required.

None of the identified outcome measures evaluated actual bimanual performance (i.e. spontaneous handling of objects requiring the use of both hands) in children with bilateral CP.

Table VIII: Focus according to ICF-CY, attribute measured and scale construction for activity-based assessments with published measurement properties in children with bilateral cerebral palsy (CP).

Outcome measure[a]	ICF-CY component[b]	ICF-CY qualifier[c]	Attribute measured	Scale construction	Published measurement properties[d]
ABILHAND-kids	A (100%)	P	Perceived Manual ability	Questionnaire	Reliability Content validity Construct validity
EDPA	A (59%)+BF	C	Unimanual Reflexes Fine motor	Based on norms	Realiability Content validity
MA	A (35%)+BF	C	Unilateral Movement quality	Criterion	Reliability Content validity
MA2	A (37%)+BF	C	Unilateral Movement quality	Criterion	Reliability Content validity Construct validity
PDMS-FMS	A (88%)+BF	C	Preferred hand Fine motor	Norm	Reliability
PDMS2-FMS	A (90%)+BF	C	Preferred hand Fine motor	Norm	Reliability Responsiveness
QUEST	A (26%)+BF	C	Unilateral Movement quality	Criterion	Reliability Content validity Construct validity

[a]BFMF=Bimanual Fine Motor Function; MACS=Manual Ability Classification System; EDPA= Erhardt Developmental Prehension Assessment; MA= Melbourne Assessment of Unilateral Upper Limb Function; MA2=Melbourne Assessment 2; PDMS=Peabody Developmental Motor Scales; FMS=Fine Motor Scale; PDMS-2=Peabody Developmental Motor Scales Second Edition; QUEST=Quality of Upper Extremity Skills Test, [b]A=activity (%=fine hand use); BF=body function, [c]C=capacity; P=performance, [d]V=validity; A=agreement (reliability); R=responsiveness

5.3 Validity and reliability of the new test BoHA

5.3.1 Item generation and content validity

The contents of the BoHA test items were discussed in the expert group until consensus was reached. Review of the 20 AHA test items revealed that five items were suitable with no changes required, nine were suitable after adaptations, while only one item was regarded as not suitable (Paper IV, Table S1, online supporting information). In addition, three potentially new BoHA items were generated. Furthermore, the appraisal of clinical relevance resulted in the exclusion of three AHA-items which showed a ceiling effect, while the original AHA items Stabilizes by weight or support and Stabilizes by grasp were merged into the modified item Stabilizes objects. Thus, the item generation resulted in 18 items for the BoHA trial version: twelve unimanual and six bimanual items. The twelve unimanual items were scored separately for the dominant (D) and the non-dominant (ND) hand.

5.3.2 **Test development, validation and aspects of reliability using Rasch analyses**

Initial analysis did not support unidimensionality of the BoHA when scorings for 171 children on 18 items were included in the Rasch analysis. The PCA for the items indicated contrasting groups of items in the dataset with the non-dominant hand and the bimanual items forming one contrast and the items for the dominant hand forming another contrast. In addition, the PCA of persons indicated contrasting groups of persons in the dataset: children with asymmetric hand use (≥20% difference between the hands; n=55) and children with more symmetric hand use (<20% difference between the hands; n=116). In addition, the DIF analysis indicated that 87% of the items functioned differently for children with asymmetric and symmetric hand use.

We therefore created two BoHA versions for the further Rasch analyses: the BoHA-S for children with symmetric hand use and the BoHA-A for children with asymmetric hand use. After removal of two misfitting items, sixteen BoHA items (11 unimanual and 5 bimanual) exhibited evidence for good internal scale validity and item (0.99 and 0.98) and person reliability (0.95 and 0.96) for the BoHA-A and BoHA-S, respectively.

By linking the BoHA measures of the two versions through anchoring of items, the measures of the respective versions were comparable, while still allowing use of separate item hierarchies. These item difficulty hierarchies indicated differences in what constitutes effective bimanual hand use for children with asymmetric or symmetric hand use, which may be useful for treatment planning (see paper IV, tables II and III). Furthermore, the person separation ratios (BoHA-A: 4.36 and BoHA-S: 5.19) indicated that both BoHA versions were able to separate bimanual performance into 6 and 7 different ability levels, respectively.

5.3.3 **Correlation to other variables**

There was a good correlation between the BoHA measures and manual abilities (MACS) (Spearman's rho: 0.74, $p \leq 0.001$). In contrast, there was low/no correlation between the BoHA measures and age (Pearson's r=0.165, p=0.035) and no correlation with gender (Spearman's rho=0.033, p=0.671).

6. DISCUSSION

6.1 Main findings

In this thesis, I have documented that fine motor capacity can be valid and reliably classified in children with CP using the BFMF. Together with the MACS classification, the BFMF provides supplementary information particularly useful for research purposes. In addition, outcome measures intended to measure change are required for evaluative purposes. Our results indicate that fine motor capacity can be assessed by the use of the MA2, while perceived manual performance can be measured by the questionnaire ABILHAND-kids. However, we found that an outcome measure of bimanual performance in children with bilateral CP was missing. Consequently, we have therefore developed a new test, the BoHA, based on modifications of the AHA. Differences in bimanual performance between children with bilateral and unilateral CP confirmed the need for a separate test for children with bilateral CP.

6.2 Validity of findings

The validity of the findings will be elaborated upon before discussing the relevance and implications of the main findings. In the following section, the possibility of systematic errors, such as selection bias or information bias in relation to the studies included in this thesis will be addressed.

Selection bias

Distorted selection of study participants from the source population can introduce selection bias and can cause effect estimates to deviate from the estimates that would have been obtained if the entire population had been included. In this thesis, there may have been four elements of selection bias: two related to the selection of children for inclusion in papers I, II and IV, and two related to language and focus in the systematic review described in paper III.

For paper I, missing BFMF or MACS classifications in the population-based CP registers in Norway and western Sweden resulted in the inclusion of 55% of the children with CP born in the defined period in Norway and 83% in western Sweden. However, since the distribution of CP subtypes and severity (GMFCS, MACS and BFMF levels) in the study population was similar to what has been reported in other population-based studies,[26,46,5] it is unlikely that this random loss of subjects with CP has led to significant selection bias.

Convenience sampling was used for papers II and IV, involving selection of the most available children. This type of sampling can lead to skewed data that may not be representative of the entire population.[70] For paper II, the distribution differed somewhat from those in population-based studies. However, the number of participants was sufficiently large to perform agreement analyses according to the COSMIN guidelines, and all CP types and functional levels were represented. Thus, it is not likely that the main results in this paper were affected by selection bias.

For paper IV, the proportion of the included children were equally distributed between the MACS levels I-III, somewhat different from a population-based study where about half of the children with bilateral CP were classified in MACS level I (55%), and 20% were classified in MACS level III.[5] However, some children classified in MACS level I may have (near) normal hand function, and maximum scores do not contribute to the calculation of item and person measures in the Rasch analysis.[77] Thus, the inclusion of equal proportions of children classified in MACS levels I-III ensured a more representative sample of the target population for the BoHA.

In the systematic review (paper III), only papers published in English were included; studies on measurement properties were excluded if assessing measurement properties of the measure in question was not the main aim of the study. This may have caused some loss of information regarding measurement properties for some of the outcome measures.

Information bias

Information bias may arise if the collected information is erroneous. In paper I, there may have been a risk of misclassification of the BFMF as fine motor performance instead of capacity. However, this is unlikely knowing that the collaborating registers were informed orally that capacity should be classified and the registers providing data for this study had originally launched the BFMF. Moreover, clinicians have traditionally focused on capacity when assessing functional abilities. For paper II, this possible cause of misclassification was eliminated by the use of a revised edition of the BFMF, which emphasizes that it is a classification of fine motor capacity.

In the systematic review, reported in paper III, two of the authors extracted data and conducted the quality assessment independently to avoid information bias. For paper IV, the test developers evaluated the suitability of the items independently and discussed discrepancies until consensus was reached. Thereafter, the final items were selected based on

Rasch modelling, ensuring internal scale validity of the included test items. However, the need to separate the children into two groups for the final Rasch analyses may have produced less precise estimates. Nevertheless, the standard errors were generally small, indicating that the number of observations used to make the estimate was sufficient.

Other methodological considerations

For paper I construct validity of the BFMF was indicated by a high correlation between the BFMF and MACS classification levels. However, this high correlation simply demonstrates that the two classification systems rank order hand function, from low to high, in the same relative order. To further investigate the validity of the differing BFMF levels, the correlation with instruments measuring capacity to grasp, hold, and manipulate should also be investigated.

For paper II, intra- and inter-reliability of the BFMF 2 was explored using video-recorded assessments of fine motor capacity as the base for the selection of each child's classification level. In addition, classifications of the BFMF levels have been determined based on information from parents, a combination of assessment and information from parents, or from medical records. Thus, there is a need for future studies assessing the reliability of the BFMF using these methods emphasizing that it is the child's best ability that is requested.

For the systematic review (paper III), the strong level of evidence found for all studies describing content validity, may indicate that the rating of this property within the COSMIN standards may not be sensitive enough to distinguish between studies. At the moment, there is an ongoing Delphi study aimed at improving this aspect within the COSMIN standards (Terwee 2015, personal communication).

For paper IV, the initial Rasch analysis indicated that we had two contrasting groups of items and persons in the dataset. To overcome this lack of unidimensionality we discussed different approaches, such as splitting the BoHA into separate scales for the dominant and the non-dominant hand, or modifying all the items into bimanual items. However, both these approaches would compromise our aims, which were to measure bimanual performance, as well as quantify a possible side difference between hands. Therefore, we chose to split the BoHA into one version for children with symmetric hand use (BoHA-S) and one for children with asymmetric hand use (BoHA-A), resulting in unidimensionality for both versions.

6.3 Assessment of different aspects of hand function in children with bilateral CP

6.3.1 The complementary contributions of the BFMF and MACS classifications

Our findings indicate that the BFMF can be used together with the MACS to classify different properties of the general construct hand function, being able to separate between five distinct severity levels in a total population of children with CP. For both classifications, the severity levels are based on the ICF five-level ordinal qualifier grading the extent and magnitude of a problem in functioning as no problem, a mild, moderate, severe, or complete problem respectively.[23,46,45] The high correlation between the BFMF and MACS demonstrates that the two classification systems rank hand function, in these broad categories, in the same relative order. Still, there are some important differences between the classification systems suggesting that they may provide complementary information regarding children's hand function if used together.

The first difference relates to the capacity and performance qualifiers described in the ICF-CY.[24] The BFMF classifies fine motor capacity according to the child's best ability to grasp, hold, and manipulate objects. Thus, capacity reflects the child's highest probable level of functioning (what the child *can* do) in a uniform or standard environment. The purpose of the MACS is to reflect manual performance by classifying the ability to handle objects (what the child *does* do) in his or her real-life environment. In other words, capacity describes functioning in a situation where one tries to eliminate the effect of the context, while performance describes functioning in the interaction between the child and the context. This indicates that the difference between capacity and performance may indicate strengths in a child's capacity to use its hands that are currently not in use and may serve as a guide as to what can be done to the environment of a child to facilitate performance.[23,24] For children with unilateral spastic CP, the need for differentiation between fine motor capacity and manual performance has been recognized clinically both for evaluation and treatment planning,[136,7] whereas little is known about the relationship between fine motor capacity and manual performance for other CP subtypes. Our finding of a systematic difference between fine motor capacity and manual performance, also for children with bilateral CP, indicate that this topic needs to be further investigated.

The second difference relates to the difference in scope of the classifications. This was indicated by the linking of meaningful concepts to the ICF-CY reflecting the more specific functions classified by the BFMF and the broader perspective classified by the MACS.

According to this, the BFMF can be said to classify simple tasks, or more "pure" motor function, placing relatively modest demands on attention, memory, and/or processing. In contrast, the MACS classifies complex tasks where various factors in the environment, as well as personal factors such as motivation, inner drive and cognition, affect the how the hands are used in daily life.[137,24]

The two classification systems may therefore be used for different grounds in intervention studies and other studies where the aim is to group children according to common characteristics. The MACS has been used for this purpose in several studies,[6,138,139] while the BFMF has mainly been used in population-based studies to describe varying levels of hand function. However, a couple of studies have indicated the need for a classification of grasping ability also in other types of studies. For example, in one study describing longitudinal development of hand function in children with unilateral CP, differences in ability to grasp and hold objects were found to be useful for describing growth curves.[13] In addition, an arbitrary classification of grasping ability was used in the test development of the Mini-AHA.[111] This suggests the need for a quick and easy classification of grasping ability, like the BFMF, that is both valid and reliable. For a classification to be feasible, it should be quick and easy to use, with clear administration guidelines, as well as meaningful and understandable descriptions of the distinctions between the classification levels.[45] So far, the BFMF has mainly been used in CP registers and population-based studies. However, the evidence of validity and reliability presented in this thesis indicates that the revised BFMF 2 may be a feasible classification that can be used in research and hopefully also in clinical practice.

However, stability over time (test-retest) for the BFMF needs further investigation to provide more evidence of the usefulness of the BFMF as a predictive instrument.[52] This has been studied as part of the SPARCLE project (Study of participation of children with cerebral palsy living in Europe), indicating that the BFMF classification levels were stable over time for the majority of the children (κ_w: > 0.75). However, the lower boundary of the 95% CI was 0.74, indicating some uncertainty about the strength of the agreement.[140] Thus, further studies are required to investigate whether the lower agreement found for the BFMF depends on variation within the children or within the raters' understanding of the BFMF classification levels.

Besides the BFMF, no valid and reliable classification of fine motor function exists for children with CP, whereas the MACS has evidence of strong measurement properties as a classification of manual performance. The House functional classification has been used in some studies to describe grip function in each hand separately. However, the House functional classification was developed for evaluative purposes,[141] and has been further modified with the aim of improving the evaluative usefulness of the measure.[142,143] The name of the House functional classification is therefore somewhat misleading, indicating that it is a functional classification while in fact it aims to be an evaluative outcome measure.

6.3.2 **Evaluation of hand function in children with bilateral CP**

Outcome measures with sound measurement properties are crucial to make decision about treatment strategies for the individual child and to evaluate the effectiveness of interventions.[144] It has been recognized that various outcome measures are required depending on which aspects of hand function we want to evaluate. Furthermore, the outcome measure in question needs to be validated for the CP subtype of interest and should provide a unidimensional interval scale to be useful for evaluative purposes.[57,78]

The systematic appraisal of outcome measures and the development of the new test BoHA have contributed to increased evidence regarding appropriate measures that can be used to evaluate various aspects of hand function in children with bilateral CP. The Abilhand-kids was developed and validated for children with all CP subtypes in 2004 and provides information regarding perceived manual performance in daily activities for children between five to sixteen years of age. The MA2 was published in 2014 and is also validated for children with CP of all subtypes. Movement quality (i.e. range of movement, accuracy, and fluency) and capacity to reach, grasp, release and manipulate (i.e. dexterity) in each hand separately can be assessed using the MA2. Using the BoHA it is possible to assess and describe how children with bilateral CP spontaneously use the hands together when handling objects requiring the use of both hands. In the BoHA test situation, the test-kit consists of toys that all require bimanual hand use. However, it is not predetermined how the child should manipulate the toys and the scorings are made from observation of the child's spontaneous handling of the toys. This makes the BoHA a measure of a child's bimanual performance, different from the MA2 assessing the child's best capacity to reach, grasp, release and manipulate on request.[10]

Another consequence of the difference in test situations between the BoHA and the MA2 relates to the age bands for which the outcome measures are appropriate. The use of a playful test situation where the children were given help when needed (reflected in the scorings), enabled testing of children with bilateral CP from the age of 18 months. In the MA2 it was found that several children below the age of 30 months were unable to complete the assessment due to difficulties attending adequately to the instructions.[145,146] The increasing focus on early interventions in children with CP highlights the need for valid and reliable outcome measures already from infancy to enable evaluation of efficacy.[124,137,143] For children with unilateral CP, an extended version of the AHA, the Mini-AHA, has been validated for children from 8 to 18 months. Similar to the AHA, the Mini-AHA evaluates effective use of the affected hand in bimanual play performance using age appropriate toys.[111] Hopefully, a corresponding outcome measure can be developed and validated also for infants and toddlers with bilateral CP.

6.3.3 Measurement properties and feasibility of the new test BoHA

6.3.3.1 Validity

The primary aim of test development is to create a valid, unidimensional measure of the underlying construct. In the development of the BoHA, we built on the theoretical construct already developed by the AHA, by changing the concept of the test into being a measure of the child's effective use of both hands together in bimanual task performance (play), as opposed to the effective use of the assisting hand.[7]

Observed differences in bimanual performance between children with bilateral and unilateral CP confirmed the slightly divergent theoretical construct underlying the BoHA and the AHA, and confirmed the need for separate measures. For example, hand role differentiation was less obvious in children with bilateral CP. The majority of the children used both hands frequently and the hand closest to the objects was commonly used, whether it was the dominant or the non-dominant hand. However, frequent use of both hands did not necessarily indicate good coordination. Actually, the coordination item was one of the more difficult items for children with bilateral CP, especially among children with symmetric hand use. In contrast, the ability to coordinate was one of the easier items for children with unilateral CP.[18] Decreased coordination in children with bilateral CP may be related to pronounced slowness of movements observed in several of these children, making bimanual performance somewhat ineffective even in children who otherwise had good abilities. Furthermore, imprecise or exaggerated movements were commonly observed in children with bilateral CP, while decreased range of motion, which is common in children with unilateral CP, was not so evident. Overall, a larger variation in bimanual hand use was observed in children with bilateral CP compared with children with unilateral CP.

The initial Rasch analysis reflected the larger variability in bimanual hand use in children with bilateral CP and indicated the need to separate the BoHA into two versions: one for children with asymmetric hand use (BoHA-A) and one for children with symmetric hand use (BoHA-S). The logical hierarchical ordering of the items in both versions confirmed the underlying construct of the BoHA and corresponded well with what we expected from clinical experience. For example for children with asymmetric hand use, all items for the non-dominant hand were more difficult than the items for the dominant hand. In contrary, for children with more symmetric hand use, the overall pattern indicated that unilateral items that were difficult for the non-dominant hand (i.e. to manipulate and grip force regulation) were also rather difficult for the dominant hand, while easy items for the dominant hand (i.e. to initiate) were similarly rather easy for the non-dominant hand. Furthermore, the absence of correlation between age and the BoHA measures suggested that the item difficulty hierarchies probably reflected steps of increasing ability rather than age-dependent development. Thus, the possibility of using BoHA for treatment planning and evaluation seems promising, but remains to be further investigated.

6.3.3.2 Reliability

Evidence of aspects of reliability of the BoHA was also demonstrated using the Rasch analysis to provide each measure and each calibration with information about its standard error (SE), indicating the precision of the measure. For the BoHA, the SEs for both versions were generally small, indicating that the number of observations used to make the estimate was sufficient even for the BoHA-A version including only 55 children.[127] Still, further analyses including a larger number of children are desired to confirm our results.

The person separation ratio was used to estimate the numbers of strata it is possible to distinguish in the sample (using three SEs to define each strata). The BoHA measures were able to separate the children in six and seven distinct difficulty strata, indicating that the test may be sensitive to detect differences among person ability measures. Furthermore, the item reliability coefficients for the BoHA-A and BoHA-S versions (0.99 and 0.98) indicated a high degree of replication of the item hierarchy with a different set of persons, and the person reliability coefficients (0.95 and 0.96) reflected a high ability of the scale to reliably rank person's relative measure location. Acceptable reliability is indicated by item and person reliability coefficients ≥0.80.[147]

Further research is needed to evaluate additional aspects of reliability (test-retest, intra- and inter-rater).

6.3.3.3 Feasibility of the BoHA

The feasibility addresses the burden of the BoHA both for the examiner and for the child and family. Some assessments take a long time to complete or might only be undertaken by examiners with certain skills.[148] Feasibility for the child being tested relates to whether or not they are able to perform the test, and also possible burden related to the physiological distress of testing both for the child and the parents.[149,150]

Similar to the AHA, the BoHA was administered in two steps. First, a semi-structured video-recorded play session was conducted in which specific toys from the AHA test kit requiring bimanual handling were used. Second, the scoring was performed by a review of the video recordings.[11] This two-step assessment using video recordings was rather time-consuming, but enabled the examiner to interact and play with the child something which is important for the feasibility of the test for the child.[7] This is similar to the AHA assessment where it has been reported that the extensive training and the rather time-consuming evaluation from the video-recordings may be a drawback.[151] To minimize the time requirements for the examiner, it was therefore considered important to maintain a play session length to a maximum of 15-20 minutes and to reduce the number of scoring items to a minimum.

According to feasibility for the child and the parents, we experienced that all children were able to perform the BoHA using the toys from the AHA test-kit, and most of the children clearly enjoyed the play session. In addition, several parents expressed that we were able to capture the child's strengths in the use of both hands together. This is similar to what has been reported from AHA users.[7,151] The rationale behind the development of the AHA test-kit was to provide a strength-based test within a playful test situation enabling testing for children from 18 months of age.[7] Focusing on the child's strengths has been regarded as very important by parents, facilitating hope and a positive focus for the growth of the child.[149]

7. CONCLUSIONS AND CLINICAL IMPLICATIONS

In this thesis, I have demonstrated that the BFMF is a classification that provides a description of fine motor capacity in a valid and reliable way for children and adolescents with CP. Fine motor capacity is defined as the child's best ability to grasp, hold and manipulate objects with each hand separately. Thus, the BFMF may be used together with the MACS, classifying children's self-initiated ability to handle objects in daily activities (manual performance), to get a more comprehensive description of children's hand function.

The BFMF can be used consistently by trained therapists, as demonstrated by the excellent intra- and inter-rater reliability. Thus, continued use of the BFMF as a classification of hand function in population based studies can be recommended. Furthermore, the use of the BFMF to group children according to grasping ability for intervention studies and longitudinal studies may be valuable. In addition, the clinical use of the BFMF may facilitate the communication between clinicians and parents, when describing the child's fine motor abilities.

Through the development and validation of the BoHA, measuring bimanual performance, I hope that hand function assessments and interventions in children with bilateral CP will be more evidence-based. In this thesis, high sensitivity of the BoHA was indicated. Thus, the ability of the BoHA to evaluate change seems promising, but needs further investigation. Children with bilateral CP spend much of their time on treatment and training. It is therefore important to provide a necessary and reliable basis for describing a child's hand function, to be able to plan and evaluate interventions, so that children avoid spending resources on inefficient treatment, or treatments that does not matter for the individual child. The use of the MA2 together with the BoHA could be useful to measure the relationship between changes in unimanual capacity and bimanual performance. In addition, the use of the ABILHAND-Kids provides a measure of overall manual ability from a parent's perspective. Thus, these three measures, when used together, may give complementary information regarding a child's hand function. Further studies of the responsiveness of these three measures are required to ensure valid measures of clinically meaningful change of hand function in children with bilateral CP.

8. FUTURE RESEARCH

During this doctoral work, new research questions have emerged. For future use, further investigation of reliability of the BFMF when classified based on information from parents, a combination of assessment and information from parents, or from medical records, is required. Furthermore, the stability of the BFMF over time needs to be better established, to determine the usefulness of the BFMF as a predictive tool in intervention and longitudinal studies. Furthermore, the correlation between the BFMF and assessments of fine motor function should be further investigated to add evidence to the construct validity of the BFMF.

The BoHA seems like a promising outcome measure to plan and evaluate efficacy of interventions, as well as for describing change in hand function in longitudinal studies. These are important areas for future research. However, before the BoHA can be used for these purposes, further studies are required to establish evidence of the responsiveness of the measure. In addition, intra- and inter-reliability need to be established, to ensure the consistency of describing bimanual performance using the BoHA. Investigation of correlation between BoHA and other assessments measuring similar or different constructs will provide further information regarding the generalizability of the BoHA results. Furthermore, it would be useful to develop outcome measures assessing bimanual performance in younger and older children, to be able to assess hand function from infants to adults, similar to the AHA.

9. REFERENCES

1. Palisano R, Rosenbaum P, Walter S, Russell D, Wood E, Galuppi B. Development and reliability of a system to classify gross motor function in children with cerebral palsy. *Dev Med Child Neurol.* 1997; **39**: 214-23.
2. Beckung E, Hagberg G. Correlation between ICIDH handicap code and Gross Motor Function Classification System in children with cerebral palsy. *Dev Med Child Neurol.* 2000; **42**: 669-73.
3. Cieza A, Brockow T, Ewert T, Amman E, Kollerits B, Chatterji S, Ustun TB, Stucki G. Linking health-status measurements to the International Classification of Functioning, Disability and Health. *J Rehabil Med.* 2002; **34**: 205-10.
4. Cieza A, Geyh S, Chatterji S, Kostanjsek N, Ustun B, Stucki G. ICF linking rules: an update based on lessons learned. *J Rehabil Med.* 2005; **37**: 212-8.
5. Arner M, Eliasson AC, Nicklasson S, Sommerstein K, Hagglund G. Hand function in cerebral palsy. Report of 367 children in a population-based longitudinal health care program. *J Hand Surg Am.* 2008; **33**: 1337-47.
6. Öhrvall AM, Eliasson AC, Lowing K, Odman P, Krumlinde-Sundholm L. Self-care and mobility skills in children with cerebral palsy, related to their manual ability and gross motor function classifications. *Dev Med Child Neurol.* 2010; **52**: 1048-55.
7. Krumlinde-Sundholm L, Eliasson AC. Development of the Assisting Hand Assessment: a Rasch-built measure intended for children with unilateral upper limb impairments. *Scand J Occup Ther.* 2003; **10**: 16-26.
8. Holmefur M, Aarts P, Hoare B, Krumlinde-Sundholm L. Test-retest and alternate forms reliability of the assisting hand assessment. *J Rehabil Med.* 2009; **41**: 886-91.
9. Holmefur M, Krumlinde-Sundholm L, Eliasson AC. Interrater and intrarater reliability of the Assisting Hand Assessment. *Am J Occup Ther.* 2007; **61**: 79-84.
10. Holmefur M, Krumlinde-Sundholm L. Psychometric properties of a revised version of the Assisting Hand Assessment (Kids-AHA 5.0). *Dev Med Child Neurol.* 2015.
11. Krumlinde-Sundholm L, Holmefur M, Kottorp A, Eliasson AC. The Assisting Hand Assessment: current evidence of validity, reliability, and responsiveness to change. *Dev Med & Child Neurol.* 2007; **49**: 259-64.
12. Wallen M. Reflections on the contribution of the Assisting Hand Assessment. *Dev Med Child Neurol.* 2015.
13. Holmefur M, Krumlinde-Sundholm L, Bergstrom J, Eliasson AC. Longitudinal development of hand function in children with unilateral cerebral palsy. *Dev Med Child Neurol.* 2010; **52**: 352-7.
14. Henderson A, Pehoski C. *Hand function in the child: Foundations for remediation.* St. Louis, Mo.: Mosby/Elsevier; 2006.
15. Aristotle. *Aristotle's on the Soul and on Memory and Recollection.* Santa Fe: Grenn Lion Press; 2001.
16. Kuhtz-Buschbeck J, Ulmer, S. Cortical control of hand function. In: Eliasson AC, Burtner, P.A. editor. *Improving hand function in children with cerebral palsy: theory, evidence and intervention.* London: Mac Keith Press; 2008. 25-42.
17. Snyder PJ, Whitaker HA. Neurologic heuristics and artistic whimsy: the cerebral cartography of Wilder Penfield. *J Hist Neurosci.* 2013; **22**: 277-91.
18. Case-Smith J, O'Brien JC. Occupational therapy for children and adolescents. St. Louis, Missouri: Elsevier; 2015.
19. Eliasson AC. Normal and impaired development of force control in precision grip. In: Henderson A, Pehoski, C. editor. *Hand function in the child: Foundations for remediation.* Second edition. St. Louis, Missouri: Mosby/Elsevier; 2006. 45-61.

20. Eliasson AC. Improving the use of hands in daily activities: aspects of the treatment of children with cerebral palsy. *Phys Occup Ther Pediatr.* 2005; **25**: 37-60.
21. Pehoski C. Cortical control of hand-object interaction. In: Henderson A, Pehoski, C. editor. *Hand function in the child: Foundations for remediation.* Second edition. St. Louis, Missouri: Mosby/Elsevier; 2006. 3-20.
22. Greaves S, Imms C, Krumlinde-Sundholm L, Dodd K, Eliasson AC. Bimanual behaviours in children aged 8-18 months: a literature review to select toys that elicit the use of two hands. *Res Dev Disabil.* 2012; **33**: 240-50.
23. World Health Organization. International classification of functioning, disability and health: ICF. Geneva: WHO; 2001.
24. World Health Organization. International classification of functioning, disability and health: children & youth version: ICF-CY. Geneva: WHO; 2007.
25. Oskoui M, Coutinho F, Dykeman J, Jette N, Pringsheim T. An update on the prevalence of cerebral palsy: a systematic review and meta-analysis. *Dev Med Child Neurol.* 2013; **55**: 509-19.
26. Andersen GL, Irgens LM, Haagaas I, Skranes JS, Meberg AE, Vik T. Cerebral palsy in Norway: prevalence, subtypes and severity. *Eur J of Paediatr Neurol.* 2008; **12**: 4-13.
27. Odding E, Roebroeck ME, Stam HJ. The epidemiology of cerebral palsy: incidence, impairments and risk factors. *Disabil Rehabil.* 2006; **28**: 183-91.
28. Sellier E, Platt MJ, Andersen GL, Krageloh-Mann I, De La Cruz J, Cans C. Decreasing prevalence in cerebral palsy: a multi-site European population-based study, 1980 to 2003. *Dev Med Child Neurol.* 2015.
29. MacLennan AH, Thompson SC, Gecz J. Cerebral palsy: causes, pathways, and the role of genetic variants. *Am J Obstet Gynecol.* 2015.
30. O'Callaghan ME, MacLennan AH, Gibson CS, McMichael GL, Haan EA, Broadbent JL, Goldwater PN, Dekker GA. Epidemiologic associations with cerebral palsy. *Obstet. Gynecol.* 2011; **118**: 576-82.
31. Himmelmann K, Ahlin K, Jacobsson B, Cans C, Thorsen P. Risk factors for cerebral palsy in children born at term. *Acta Obstet Gynecol Scand.* 2011; **90**: 1070-81.
32. McIntyre S, Taitz D, Keogh J, Goldsmith S, Badawi N, Blair E. A systematic review of risk factors for cerebral palsy in children born at term in developed countries. *Dev Med Child Neurol.* 2013; **55**: 499-508.
33. Schifrin BS, Longo LD. William John Little and cerebral palsy. A reappraisal. *Eur J Obstet Gynecol Reprod Biol.* 2000; **90**: 139-44.
34. Morris C. Definition and classification of cerebral palsy: a historical perspective. *Dev Med Child Neurol. Supplement.* 2007; **109**: 3-7.
35. Rosenbaum P, Paneth N, Leviton A, Goldstein M, Bax M, Damiano D, Dan B, Jacobsson B. A report: the definition and classification of cerebral palsy April 2006. *Dev Med Child Neurol. Supplement.* 2007; **109**: 8-14.
36. Eliasson AC, Burtner PA. Introduction. In: Eliasson AC, Burtner, P.A. editor. Improving hand function in cerebral palsy. London: Mac Keith Press; 2008.
37. Surveillance of cerebral palsy in Europe: a collaboration of cerebral palsy surveys and registers. Surveillance of Cerebral Palsy in Europe (SCPE). *Dev Med Child Neurol.* 2000; **42**: 816-24.
38. Himmelmann K, Beckung E, Hagberg G, Uvebrant P. Gross and fine motor function and accompanying impairments in cerebral palsy. *Dev Med Child Neurol.* 2006; **48**: 417-23.
39. Himmelmann K, Uvebrant P. The panorama of cerebral palsy in Sweden. XI. Changing patterns in the birth-year period 2003-2006. *Acta Paediatr.* 2014.
40. Krageloh-Mann I, Cans C. Cerebral palsy update. *Brain Dev.* 2009; **31**: 537-44.

41. Mutlu A, Kara OK, Gunel MK, Karahan S, Livanelioglu A. Agreement between parents and clinicians for the motor functional classification systems of children with cerebral palsy. *Disabil Rehabil.* 2011; **33**: 927-32.
42. Rosenbaum PL, Walter SD, Hanna SE, Palisano RJ, Russell DJ, Raina P, Wood E, Bartlett DJ, Galuppi BE. Prognosis for gross motor function in cerebral palsy: creation of motor development curves. *JAMA.* 2002; **288**: 1357-63.
43. Palisano RJ, Cameron D, Rosenbaum PL, Walter SD, Russell D. Stability of the gross motor function classification system. *Dev Med Child Neurol.* 2006; **48**: 424-8.
44. Palisano R RP, Bartlett D, Livingston M. Gross Motor Function Classification System Expanded and Revised. 2011
45. Eliasson AC, Krumlinde-Sundholm L, Rosblad B, Beckung E, Arner M, Ohrvall AM, Rosenbaum P. The Manual Ability Classification System (MACS) for children with cerebral palsy: scale development and evidence of validity and reliability. *Dev Med Child Neurol.* 2006; **48**: 549-54.
46. Beckung E, Hagberg G. Neuroimpairments, activity limitations, and participation restrictions in children with cerebral palsy. *Dev Med Child Neurol.* 2002; **44**: 309-16.
47. Hidecker MJ, Paneth N, Rosenbaum PL, Kent RD, Lillie J, Eulenberg JB, Chester K, Jr., Johnson B, Michalsen L, Evatt M, Taylor K. Developing and validating the Communication Function Classification System for individuals with cerebral palsy. *Dev Med Child Neurol.* 2011; **53**: 704-10.
48. Sellers D, Mandy A, Pennington L, Hankins M, Morris C. Development and reliability of a system to classify the eating and drinking ability of people with cerebral palsy. *Dev Med Child Neurol.* 2014; **56**: 245-51.
49. Carnahan KD, Arner M, Hagglund G. Association between gross motor function (GMFCS) and manual ability (MACS) in children with cerebral palsy. A population-based study of 359 children. *BMC Musculoskelet Disord.* 2007; **8**: 50.
50. Gunel MK, Mutlu A, Tarsuslu T, Livanelioglu A. Relationship among the Manual Ability Classification System (MACS), the Gross Motor Function Classification System (GMFCS), and the functional status (WeeFIM) in children with spastic cerebral palsy. *Eur J Pediatr.* 2009; **168**: 477-85.
51. Jeevanantham D, Dyszuk E, Bartlett D. The Manual Ability Classification System: A Scoping Review. *Pediatr Phys Ther.* 2015; **27**: 236-41.
52. Öhrvall AM, Krumlinde-Sundholm L, Eliasson AC. The stability of the Manual Ability Classification System over time. *Dev Med Child Neurol.* 2014; **56**: 185-9.
53. Cans C, Dolk H, Platt MJ, Colver A, Prasauskiene A, Krageloh-Mann I. Recommendations from the SCPE collaborative group for defining and classifying cerebral palsy. *Dev Med Child Neurol. Supplement.* 2007; **109**: 35-8.
54. Colver A. Study protocol: SPARCLE - a multi-centre European study of the relationship of environment to participation and quality of life in children with cerebral palsy. *BMC Public Health.* 2006; **6**: 105.
55. Lien E, Andersen GL, Bao Y, Gordish-Dressman H, Skranes JS, Vik T, Blackman JA. Apolipoprotein E polymorphisms and severity of cerebral palsy: a cross-sectional study in 255 children in Norway. *Dev Med Child Neurol.* 2013; **55**: 372-7.
56. Randall M, Harvey A, Imms C, Reid S, Lee KJ, Reddihough D. Reliable classification of functional profiles and movement disorders of children with cerebral palsy. *Phys Occup Ther Pediatr.* 2013; **33**: 342-52.
57. Krumlinde-Sundholm L. Choosing and using assessments of hand function. In: Eliasson AC, Burtner, P.A. editor. *Improving hand function in children with cerebral palsy: theory, evidence and intervention.* London: Mac Keith Press; 2008. 176-97.

58. van Meeteren J, Nieuwenhuijsen C, de Grund A, Stam HJ, Roebroeck ME. Using the manual ability classification system in young adults with cerebral palsy and normal intelligence. *Disabil. Rehabil.* 2010; **32**: 1885-93.
59. Majnemer A, Mazer B. New directions in the outcome evaluation of children with cerebral palsy. *Semin Pediatr Neurol.* 2004; **11**: 11-7.
60. Vargus-Adams J. Understanding function and other outcomes in cerebral palsy. *Phys Med Rehabil Clin N Am.* 2009; **20**: 567-75.
61. Majnemer A. Benefits of using outcome measures on pediatric rehabilitation. *Phys Occup Ther Pediatr.* 2010; **30**: 165-7.
62. Rosenbaum P. Screening tests and standardized assessments used to identify and characterize developmental delays. *Semin Pediatr Neurol.* 1998; **5**: 27-32.
63. Dekker J, Dallmeijer AJ, Lankhorst GJ. Clinimetrics in rehabilitation medicine: current issues in developing and applying measurement instruments 1. *J Rehabil Med.* 2005; **37**: 193-201.
64. Gilmore R, Sakzewski L, Boyd R. Upper limb activity measures for 5- to 16-year-old children with congenital hemiplegia: a systematic review. *Dev Med Child Neurol.* 2009.
65. Greaves S, Imms C, Dodd K, Krumlinde-Sundholm L. Assessing bimanual performance in young children with hemiplegic cerebral palsy: a systematic review. *Dev Med Child Neurol.* 2010; **52**: 413-21.
66. Klingels K, Jaspers E, Van de Winckel A, De Cock P, Molenaers G, Feys H. A systematic review of arm activity measures for children with hemiplegic cerebral palsy. *Clin Rehabil.* 2010; **24**: 887-900.
67. Lemmens RJ, Timmermans AA, Janssen-Potten YJ, Smeets RJ, Seelen HA. Valid and reliable instruments for arm-hand assessment at ICF activity level in persons with hemiplegia: a systematic review. *BMC Neurol.* 2012; **12**: 21.
68. Krumlinde-Sundholm L, Ek L, Eliasson AC. What assessments evaluate use of hands in infants? A literature review. *Dev Med Child Neurol.* 2015; **57 Suppl 2**: 37-41.
69. Mokkink LB, Terwee CB, Knol DL, Stratford PW, Alonso J, Patrick DL, Bouter LM, de Vet HC. The COSMIN checklist for evaluating the methodological quality of studies on measurement properties: a clarification of its content. *BMC Med Res Methodol.* 2010; **10**: 22.
70. Portney LG, Watkins MP. *Foundations of clinical research: applications to practice (3rd ed).* New Jersey: Pearson Education; 2009.
71. Mokkink LB, Terwee CB, Knol DL, Stratford PW, Alonso J, Patrick DL, Bouter LM, de Vet HC. Protocol of the COSMIN study: COnsensus-based Standards for the selection of health Measurement INstruments. *BMC Med Res Methodol.* 2006; **6**: 2.
72. Öhrvall AM, Krumlinde-Sundholm L, Eliasson AC. Exploration of the relationship between the Manual Ability Classification System and hand-function measures of capacity and performance. *Disabil Rehabil.* 2013; **35**: 913-8.
73. Law M, McDermid J. *Evidence-based rehabilitation : a guide to practice.* Thorofare, N.J: Slack; 2008.
74. Terwee CB, Bot SD, de Boer MR, van der Windt DA, Knol DL, Dekker J, Bouter LM, de Vet HC. Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol.* 2007; **60**: 34-42.
75. Randall M, Imms C, Carey LM, Pallant JF. Rasch analysis of The Melbourne Assessment of Unilateral Upper Limb Function. *Dev Med Child Neurol.* 2014.
76. DeMatteo C, Law M, Russell D, Pollock N, Rosenbaum P, Walter S. *QUEST: Quality of upper extremity skills test manual.* Hamilton, Ontario: Chedokee-McMaster Hospitals; 1992.
77. Bond TG, Fox CM. *Applying the Rasch model: fundamental measurement in the human sciences.* New York: Taylor & Francis Group, LLC; 2007.

78. Grimby G, Tennant A, Tesio L. The use of raw scores from ordinal scales: time to end malpractice? *J Rehabil Med.* 2012; **44**: 97-8.
79. Froslev-Friis C, Dunkhase-Heinl U, Andersen JD, Stausbol-Gron B, Hansen AV, Garne E. Epidemiology of cerebral palsy in Southern Denmark. *Dan Med J.* 2015; **62**.
80. Himmelmann K, Uvebrant P. Function and neuroimaging in cerebral palsy: a population-based study. *Dev Med Child Neurol.* 2011; **53**: 516-21.
81. Chounti A, Hagglund G, Wagner P, Westbom L. Sex differences in cerebral palsy incidence and functional ability: a total population study. *Acta Paediatr.* 2013; **102**: 712-7.
82. Shevell MI, Dagenais L, Hall N. The relationship of cerebral palsy subtype and functional motor impairment: a population-based study. *Dev Med Child Neurol.* 2009; **51**: 872-7.
83. Levin MF, Sveistrup, H. Postural control for reaching and hand skills. In: Eliasson AC, Burtner, P.A. editor. *Improving hand function in children with cerebral palsy: theory, evidence and intervention.* London: Mac Keith Press; 2008. 109-23.
84. Valvano J, Newell KM. Practice of a precision isometric grip-force task by children with spastic cerebral palsy. *Dev Med Child Neurol.* 1998; **40**: 464-73.
85. Park ES, Sim EG, Rha DW. Effect of upper limb deformities on gross motor and upper limb functions in children with spastic cerebral palsy. *Res Dev Disabil.* 2011; **32**: 2389-97.
86. Gao F, Mei X, Chen AC. Delayed finger tapping and cognitive responses in preterm-born male teenagers with mild spastic diplegia. *Pediatr Neurol.* 2015; **52**: 206-13.
87. Mawase F, Bar-Haim S, Karniel A. Lack of predictive control in lifting series of virtual objects by individuals with diplegic cerebral palsy. *IEEE Trans Neural Syst Rehabil Eng.* 2011; **19**: 686-95.
88. Blank R, Hermsdorfer J. Basic motor capacity in relation to object manipulation and general manual ability in young children with spastic cerebral palsy. *Neurosci Lett.* 2009; **450**: 65-9.
89. Sanger TD. Arm trajectories in dyskinetic cerebral palsy have increased random variability. *J Child Neurol.* 2006; **21**: 551-7.
90. Sanger TD, Kaiser J, Placek B. Reaching movements in childhood dystonia contain signal-dependent noise. *J. Child Neurol.* 2005; **20**: 489-96.
91. 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. *Dev Med Child Neurol.* 2015.
92. Ahsgren I, Baldwin I, Goetzinger-Falk C, Erikson A, Flodmark O, Gillberg C. Ataxia, autism, and the cerebellum: a clinical study of 32 individuals with congenital ataxia. *Dev Med Child Neurol.* 2005; **47**: 193-8.
93. Jethwa A, Mink J, Macarthur C, Knights S, Fehlings T, Fehlings D. Development of the Hypertonia Assessment Tool (HAT): a discriminative tool for hypertonia in children. *Dev Med Child Neurol.* 2010; **52**: e83-7.
94. Sanger TD, Delgado MR, Gaebler-Spira D, Hallett M, Mink JW. Classification and definition of disorders causing hypertonia in childhood. *Pediatrics.* 2003; **111**: e89-97.
95. Sanger TD, Kukke SN. Abnormalities of tactile sensory function in children with dystonic and diplegic cerebral palsy. *J. Child Neurol.* 2007; **22**: 289-93.
96. Wingert JR, Burton H, Sinclair RJ, Brunstrom JE, Damiano DL. Tactile sensory abilities in cerebral palsy: deficits in roughness and object discrimination. *Dev Med Child Neurol.* 2008; **50**: 832-8.
97. Lesny I, Stehlik A, Tomasek J, Tomankova A, Havlicek I. Sensory disorders in cerebral palsy: two-point discrimination. *Dev Med Child Neurol.* 1993; **35**: 402-5.

98. Himmelmann K, McManus V, Hagberg G, Uvebrant P, Krageloh-Mann I, Cans C. Dyskinetic cerebral palsy in Europe: trends in prevalence and severity. *Arch Dis Child.* 2009; **94**: 921-6.
99. Dufresne D, Dagenais L, Shevell MI. Spectrum of visual disorders in a population-based cerebral palsy cohort. *Pediatr Neurol.* 2014; **50**: 324-8.
100. Sigurdardottir S, Eiriksdottir A, Gunnarsdottir E, Meintema M, Arnadottir U, Vik T. Cognitive profile in young Icelandic children with cerebral palsy. *Dev Med Child Neurol.* 2008; **50**: 357-62.
101. Novak I, Hines M, Goldsmith S, Barclay R. Clinical prognostic messages from a systematic review on cerebral palsy. *Pediatrics.* 2012; **130**: e1285-312.
102. Valvano J, Long T. Neurodevelopmental treatment a review of the writings of the Bobaths *Pediatr Phys Ther.* 1991; **3**: 125-9.
103. Novak I, McIntyre S, Morgan C, Campbell L, Dark L, Morton N, Stumbles E, Wilson SA, Goldsmith S. A systematic review of interventions for children with cerebral palsy: state of the evidence. *Dev Med Child Neurol.* 2013; **55**: 885-910.
104. Law M. DJ, Pollock N., King G., Rosenbaum P., Russell D., Palisano R., Harris S., Armstrong R., Watt J. Family-Centred Functional Therapy for Children with Cerebral Palsy. *Phys Occup Ther Pediatr.* 1998; **18**: 83-102.
105. Hanna SE, Law MC, Rosenbaum PL, King GA, Walter SD, Pollock N, Russell DJ. Development of hand function among children with cerebral palsy: growth curve analysis for ages 16 to 70 months. *Dev Med Child Neurol.* 2003; **45**: 448-55.
106. Eliasson AC, Forssberg H, Hung YC, Gordon AM. Development of hand function and precision grip control in individuals with cerebral palsy: a 13-year follow-up study. *Pediatrics.* 2006; **118**: e1226-36.
107. Blank R, Kluger G. Changes in elementary finger-hand functions over time in preschool children with spastic cerebral palsy. *Neurosci Lett.* 2009; **455**: 30-5.
108. Folio MR, & Fewell, R.E. *Peabody Developmental Motor Scales and Activity Cards.* Chicago: Riverside; 1983.
109. Taylor N, Sand PL, Jebsen RH. Evaluation of hand function in children. *Arch Phys Med Rehabil.* 1973; **54**: 129-35.
110. Eliasson AC, Gordon AM, Forssberg H. Basic co-ordination of manipulative forces of children with cerebral palsy. *Dev Med Child Neurol.* 1991; **33**: 661-70.
111. Greaves S, Imms C, Dodd K, Krumlinde-Sundholm L. Development of the Mini-Assisting Hand Assessment: evidence for content and internal scale validity. *Dev Med Child Neurol.* 2013.
112. Gordon AM, Bleyenheuft Y, Steenbergen B. Pathophysiology of impaired hand function in children with unilateral cerebral palsy. *Dev Med Child Neurol.* 2013; **4**: 32-7.
113. Bleyenheuft Y, Gordon AM. Precision grip control, sensory impairments and their interactions in children with hemiplegic cerebral palsy: a systematic review. *Res Dev Disabil.* 2013; **34**: 3014-28.
114. Eliasson AC, Gordon AM, Forssberg H. Impaired anticipatory control of isometric forces during grasping by children with cerebral palsy. *Dev Med Child Neurol.* 1992; **34**: 216-25.
115. Hung YC, Charles J, Gordon AM. Bimanual coordination during a goal-directed task in children with hemiplegic cerebral palsy. *Dev Med Child Neurol.* 2004; **46**: 746-53.
116. Vaz DV, Cotta Mancini M, Fonseca ST, Vieira DS, de Melo Pertence AE. Muscle stiffness and strength and their relation to hand function in children with hemiplegic cerebral palsy. *Dev Med Child Neurol.* 2006; **48**: 728-33.

117. Braendvik SM, Elvrum AK, Vereijken B, Roeleveld K. Relationship between neuromuscular body functions and upper extremity activity in children with cerebral palsy. *Dev Med Child Neurol.* 2010; **52**: e29-34.
118. Braendvik SM, Elvrum AK, Vereijken B, Roeleveld K. Involuntary and voluntary muscle activation in children with unilateral cerebral palsy--relationship to upper limb activity. *Eur J Paediatr Neurol.* 2013; **17**: 274-9.
119. Eliasson AC, Gordon AM. Impaired force coordination during object release in children with hemiplegic cerebral palsy. *Dev Med Child Neurol.* 2000; **42**: 228-34.
120. Auld ML, Boyd R, Moseley GL, Ware R, Johnston LM. Tactile function in children with unilateral cerebral palsy compared to typically developing children. *Disabil Rehabil.* 2012; **34**: 1488-94.
121. Bodimeade HL, Whittingham K, Lloyd O, Boyd RN. Executive function in children and adolescents with unilateral cerebral palsy. *Dev Med Child Neurol.* 2013; **55**: 926-33.
122. Kakooza-Mwesige A, Forssberg H, Eliasson AC, Tumwine JK. Cerebral palsy in children in Kampala, Uganda: clinical subtypes, motor function and co-morbidities. *BMC Res Notes.* 2015; **8**: 015-1125.
123. Sellier E, Uldall P, Calado E, Sigurdardottir S, Torrioli MG, Platt MJ, Cans C. Epilepsy and cerebral palsy: characteristics and trends in children born in 1976-1998. *Eur J Paediatr Neurol.* 2012; **16**: 48-55.
124. Eliasson AC, Sjostrand L, Ek L, Krumlinde-Sundholm L, Tedroff K. Efficacy of baby-CIMT: study protocol for a randomised controlled trial on infants below age 12 months, with clinical signs of unilateral CP. *BMC Pediatr.* 2014; **14**: 1471-2431.
125. Fedrizzi E, Pagliano E, Andreucci E, Oleari G. Hand function in children with hemiplegic cerebral palsy: prospective follow-up and functional outcome in adolescence. *Dev Med Child Neurol.* 2003; **45**: 85-91.
126. Streiner DL, Norman GR. *Health measurement scales: a practical guide to their development and use.* Oxford: Oxford University Press; 2015.
127. Linacre JM. Winsteps ® *Rasch measurement computer program User's Guide*. Beaverton, Oregon: Winsteps.com; 2014.
128. Terwee CB, Mokkink LB, Knol DL, Ostelo RW, Bouter LM, de Vet HC. Rating the methodological quality in systematic reviews of studies on measurement properties: a scoring system for the COSMIN checklist. *Qual Life Res.* 2012; **21**: 651-7.
129. van Tulder M, Furlan A, Bombardier C, Bouter L. Updated method guidelines for systematic reviews in the cochrane collaboration back review group. *Spine (Phila Pa 1976).* 2003; **28**: 1290-9.
130. Arnould C, Penta M, Renders A, Thonnard JL. ABILHAND-Kids: a measure of manual ability in children with cerebral palsy. *Neurology.* 2004; **63**: 1045-52.
131. Erhardt RP. *Developmental Hand Dysfunction: Theory-Assessment-Treatment.* Laurel MD, Ramsco Publishing Company; 1982.
132. Randall M, Johnson LM, Reddihough DS. *The Melbourne Assessment of Unilateral Upper Limb Function: test administration manual.* Melbourne: Royal Children's Hospital; 1999.
133. Randall M, Johnson, L., Reddihough, D. The Melbourne Assessment 2: A test of Unilateral Upper Limb Function
134. Folio MR, & Fewell, R.E. *Peabody Developmental Motor Scales Second Edition.* Austin, Texas: PRO-ED; 2000.
135. Wallen M. Clinical and research considerations in using the Melbourne Assessment 2. *Dev Med Child Neurol.* 2014; **56**: 608-9.
136. Facchin P, Rosa-Rizzotto M, Visona Dalla Pozza L, Turconi AC, Pagliano E, Signorini S, Tornetta L, Trabacca A, Fedrizzi E. Multisite trial comparing the efficacy of

constraint-induced movement therapy with that of bimanual intensive training in children with hemiplegic cerebral palsy: postintervention results. *Am J Phys Med Rehabil.* 2011; **90**: 539-53.
137. Wulf G, Shea CH. Principles derived from the study of simple skills do not generalize to complex skill learning. *Psychon Bull Rev.* 2002; **9**: 185-211.
138. Kuijper MA, van der Wilden GJ, Ketelaar M, Gorter JW. Manual ability classification system for children with cerebral palsy in a school setting and its relationship to home self-care activities. *Am J Occup Ther.* 2010; **64**: 614-20.
139. Gong HS, Chung CY, Park MS, Shin HI, Chung MS, Baek GH. Functional outcomes after upper extremity surgery for cerebral palsy: comparison of high and low manual ability classification system levels. *J Hand Surg Am.* 2010; **35**: 277-83.
140. Nystrand M, Beckung E, Dickinson H, Colver A. Stability of motor function and associated impairments between childhood and adolescence in young people with cerebral palsy in Europe. *Dev Med Child Neurol.* 2014; **56**: 833-8.
141. House JH, Gwathmey FW, Fidler MO. A dynamic approach to the thumb-in palm deformity in cerebral palsy. *J. Bone Joint Surg. Am.* 1981; **63**: 216-25.
142. Koman LA, Williams RM, Evans PJ, Richardson R, Naughton MJ, Passmore L, Smith BP. Quantification of upper extremity function and range of motion in children with cerebral palsy. *Dev Med Child Neurol.* 2008; **50**: 910-7.
143. Geerdink Y, Lindeboom R, de Wolf S, Steenbergen B, Geurts AC, Aarts P. Assessment of upper limb capacity in children with unilateral cerebral palsy: construct validity of a Rasch-reduced Modified House Classification. *Dev Med Child Neurol.* 2014; **56**: 580-6.
144. Bower E. Using the Assisting Hand Assessment and the Mini-AHA for clinical evaluation and further research and development. *Dev Med Child Neurol.* 2013; **55**: 977-8.
145. Randall M, Imms C, Carey L. Establishing validity of a modified Melbourne Assessment for children ages 2 to 4 years. *Am J Occup Ther.* 2008; **62**: 373-83.
146. Randall M, Imms C, Carey L. Further evidence of validity of the Modified Melbourne Assessment for neurologically impaired children aged 2 to 4 years. *Dev Med Child Neurol.* 2012; **54**: 424-8.
147. Fisher Jr W. Rating Scale Instrument Quality Criteria. *Rasch Measurement Transactions.* 2007; **21**: 1095.
148. Kucukdeveci AA, Tennant A, Grimby G, Franchignoni F. Strategies for assessment and outcome measurement in physical and rehabilitation medicine: an educational review. *J Rehabil Med.* 2011; **43**: 661-72.
149. Klein S, Wynn K, Ray L, Demeriez L, LaBerge P, Pei J, St Pierre C. Information sharing during diagnostic assessments: what is relevant for parents? *Phys Occup Ther Pediatr.* 2011; **31**: 120-32.
150. de Vet HCW, Terwee CB, Mokkink LB, Knol DL. *Measurement in Medicine : A Practical Guide.* Cambridge: Cambridge University Press; 2011.
151. Gordon AM. Measuring 'activity limitation' in individuals with unilateral upper extremity impairments. *Dev Med Child Neurol.* 2007; **49**: 245.

Paper I

Elvrum, Ann-Kristin; Andersen, Guro Lillemoen; Himmelmann, Kate; Beckung, Eva; Öhrvall, Ann-Marie; Lydersen, Stian; Vik, Torstein. Bimanual Fine Motor Function (BFMF) classification in children with cerebral palsy: Aspects of construct and content validity. Physical & Occupational Therapy in Pediatrics 2016 ;Volum 36.(1) s. 1-16. Not included due to copyright, available at http://doi.org/10.3109/01942638.2014.975314

Is not included due to copyright

Paper II

Bimanual capacity of children with cerebral palsy: Intra- and inter-rater reliability of a revised edition of the Bimanual Fine Motor Function classification

Ann-Kristin G. Elvrum,[1, 2] Eva Beckung,[3] Rannei Sæther,[2,4] Stian Lydersen,[5] Torstein Vik[2] & Kate Himmelmann[6]

[1]Department of Clinical Services, St. Olavs University Hospital, Trondheim, Norway,

[2]Department of Laboratory Medicine, Children's and Women's Health, Faculty of Medicine, Norwegian University of Science and Technology, Trondheim, Norway,

[3]Department of Physiotherapy, Institute of Neuroscience and Physiology, Sahlgrenska Academy at the University of Gothenburg, Gothenburg, Sweden,

[4]Department of Paediatrics, St. Olavs Hospital, Trondheim University Hospital, Norway

[5]Regional Centre for Child and Youth Mental Health and Child Welfare, Faculty of Medicine, Norwegian University of Science and Technology, Trondheim, Norway

[6]Department of Pediatrics, Institute of Clinical Sciences, Queen Silvia Children's Hospital, Sahlgrenska Academy at the University of Gothenburg, Gothenburg, Sweden

Address correspondence to:

Ann-Kristin G. Elvrum, Department of Clinical Services, St. Olavs University Hospital, Postboks 3250 Sluppen, NO 7006 Trondheim, Norway (E-mail: ann-kristin.elvrum@ntnu.no)

ABSTRACT. *Aims:* To develop a revised edition of the Bimanual Fine Motor Function (BFMF), as a classification of fine motor capacity in children with cerebral palsy (CP), and to explore the intra- and inter-rater reliability of this edition. *Methods:* The content of the original BFMF was discussed by an expert panel, resulting in a revised edition comprising the original description of the classification levels, but in addition including figures with specific explanatory text. Using this edition, four professionals classified fine motor function of 79 children (3-17 years; 45 boys) with all CP subtypes (Manual Ability Classification levels I-V). Intra- and inter-rater reliability was assessed using overall intra-class correlation coefficient (ICC), and Cohen's quadratic weighted kappa. *Results:* The overall ICC was 0.86, and Cohen's weighted kappa indicated high intra-rater (κ_w: >0.90) and inter-rater (κ_w: > 0.85) reliability. *Conclusions:* The new BFMF edition had high intra- and inter-rater reliability. The classification levels could be determined from short video recordings (< 5 minutes), using the figures and precise descriptions of the fine motor function levels included in the BFMF 2. Thus, the BFMF 2 may be a feasible and useful classification of fine motor capacity both in research and in clinical practice.

KEYWORDS: Bimanual fine motor function, cerebral palsy, classification, capacity, hand function, reliability

Children with cerebral palsy (CP) display varying degrees of limitations in functional use of the hands, ranging from difficulties only with in-hand manipulation to more severe impairments that make it impossible to even grasp or hold (Arner et al., 2008; Krumlinde-Sundholm et al., 2007). Depending on the severity of the limitations, children experience varying degrees of difficulties in the performance of daily activities in self-care, school, play and leisure (Öhrvall et al., 2010). Functional classifications of hand function provide common terminology and descriptions that can improve the understanding of the diversity in manual impairments in children with CP. They can also assist in identification of best practice for children who perform at different levels. This is crucial for prognostic counselling, and for appropriate planning and evaluation of interventions.

Bimanual Fine Motor Function (BFMF) is a classification of the hand function in children with CP on a five-level scale, whereby level I describes the best and level V the most limited function (Beckung & Hagberg, 2002). This classification system was first published in 2002, and has since been used in a number of epidemiological studies and in several CP registers included in the common data base established by the Surveillance of Cerebral Palsy in Europe (SCPE) (Andersen et al., 2008; Beckung & Hagberg, 2002; Colver, 2006; Himmelmann et al., 2006; Lien et al., 2013). The stability of BFMF over time, has been studied as part of the SPARCLE project (Study of participation of children with cerebral palsy living in Europe), indicating that the classification level was stable over time for the majority of the children (Nystrand et al., 2014). Nonetheless, the documentation of validity and reliability of BFMF has been scarce.

In a recent study, the content and construct validity of the BFMF was explored (Elvrum et al., 2014). The conclusion was that the BFMF could be a useful classification for describing fine motor capacity (what the child *can* do), adding complementary information regarding children's hand function, when used together with the Manual Ability Classification System (MACS) (Eliasson et al., 2006). The latter classification is extensively used both in clinical practice and in research, providing valid and reliable information of manual performance (what the child *does* do) (Eliasson et al., 2006; Jeevanantham et al., 2015). Moreover, the BFMF also offers the possibility to classify the capacity of the two hands separately, in contrast to MACS classifying the performance of both hands together. However, even though the BFMF was developed to classify the construct fine motor capacity (i.e. capacity to grasp, hold and manipulate), this was not specified in the original description. (Beckung & Hagberg, 2002). Furthermore, the descriptions used to differentiate between the five levels could have been more clearly explained to facilitate the use of the classification.

Thus, there is a need to develop a revised edition of the BFMF, as a classification of fine motor capacity, including a clear description of the constructs needed to separate the five levels (Elvrum et al., 2014). Furthermore, there is a need to document intra- and inter-rater reliability of the BFMF using appropriate sample sizes. So far only one study has investigated inter-rater reliability of the BFMF in children with CP (Randall et al., 2013). This study indicated high inter-rater reliability, however including only 20 children.

The aim of this paper was to present the revised edition of the BFMF, emphasizing its purpose as a classification of fine motor capacity, and to describe the results of the evaluation of intra- and inter-rater reliability of the new edition in children with CP.

METHODS

The first step was to develop an easy-to-use, revised edition of the BFMF (hereafter named BFMF 2), adding more precise descriptions of the levels (I-V) and sublevels (a) and (b) of fine motor function to facilitate the understanding and applicability of the BFMF. The wording and content of the different levels of the original BFMF were discussed and evaluated by one pediatrician, one pediatric neurologist, one physiotherapist and two occupational therapists. The developer of the original BFMF (EB) was member of this expert panel. It was agreed that the original BFMF did not clearly define the four main terms (‘without restriction’, ‘restriction in advanced fine motor skills’, ‘can grasp and hold, no in-hand manipulation’ and ‘may hold’), which were used to separate between the five classification BFMF levels. To be able to describe these terms more precisely, video recordings of children performing at different levels of fine motor function were scrutinized and observable actions were described. In addition, explanatory figures for each of the five classification levels were constructed to describe the various combinations of limited hand function in each hand. Thus, the revised BFMF 2 consisted of explanatory figures and text that should be used together with the original description of the five BFMF classification levels and sublevels (a) and (b) (see Appendix S1, online supporting information).

The second step, was to investigate the intra- and inter-rater reliability of the BFMF 2. For this purpose we aimed to recruit at least 50 participants, in line with the COSMIN guidelines (Mokkink et al., 2010). A convenience sample of 79 children, with all CP subtypes, and at all MACS and Gross Motor Function Classification System (GMFCS) (Palisano R, 2007) levels, was recruited at two study sites; the Regional Rehabilitation Centre, Queen Silvia Children’s Hospital, Göteborg, Sweden and St Olav Hospital in Trondheim, Norway. The children were 3-17 years old (mean age 8 years and 7 months, SD 3 years and 2 months) and 45 (57%) were boys. Nineteen children (24%) had unilateral spastic CP, 42 (53%) had bilateral spastic CP, 15 (19%) had dyskinetic and three children (4%) had the ataxic CP subtype. The CP classification of the Surveillance of Cerebral Palsy in Europe was applied (SCPE, 2001). The distribution of the MACS and GMFCS levels by CP type is shown in Table 1.

After oral and written information to the parents, and adapted information to the children, written informed consent was obtained. Thirty-three children were exclusively recruited to the current study, while 46 participated in other studies where video assessment of hand function was part of the study protocol.

Ethical approval for this study was granted by the Regional Ethical Committee (REK) for Medical Research in Mid-Norway (ref. 2012/152) and by the Regional Ethical Review Board, University of Gothenburg, Sweden (683-12, T377-15).

Procedure

Assessment of each child's fine motor capacity was video recorded. The child was asked to perform tasks with each hand separately, involving grasping and holding of three objects of different size (pencil, cereal/raison/small piece of chocolate and cube), as well as in-hand-manipulation of the cube. First the child was requested to grasp the object from the table. If this was not possible, the task was adapted and the child was asked to grasp the object from the examiners hand. If this was too difficult, the object was placed in the child's hand to assess whether or not the child was able to hold. Furthermore, in-hand manipulation was assessed by asking the child to grasp the cube, hold it in one hand, and turn it over and around in all directions using small finger movements. If the child was not able to do this, he or she was asked to turn the cube over and around on the table surface. This assessment took between three to five minutes to perform.

Four raters classified the children's fine motor function (n=79) from video recordings applying the BFMF 2. Two of the raters (KH and EB), were situated at the Sahlgrenska Academy at the University of Gothenburg, Gothenburg, Sweden. The two others (AKGE and RS), were situated at the Norwegian University of Science and Technology, Trondheim, Norway. All the raters were experienced clinicians. Three of the raters (KH, EB and AKGE) developed the new edition BFMF 2. In addition a fourth rater (RS) was included who was not familiar with the BFMF. To familiarize herself with the BFMF 2, she classified eight children, not included in the study and discussed her classifications with the other raters before the study started.

Two sets of the video recordings were randomly ordered, using a random order generator, where each child received a different registration number for each of the two sets. Each site received both sets of the video recordings, and the four raters independently classified fine motor function according to the BFMF 2 for all the children, at two different time-points (test replication one and two) at least two weeks apart. A key linking the two sets of registration numbers, and the individual classifications done by the raters, were stored by a third party (TV) until all assessments had been completed.

Statistics

The overall intraclass correlation coefficient (ICC) was calculated using a mixed model with the rating level as dependent variable, the test replications (one and two) as categorical covariate (also known as a fixed factor), and the children and the raters as crossed random factors. With this analysis, it could determine whether certain raters tended to give consistently higher scores than other raters, quantify variance components due to variation between children, between raters, and residual variance, and estimate the ICC.

The intra- and inter-rater reliability of the BFMF levels I-V, was analyzed in terms of absolute agreement and agreement by Cohen’s quadratic weighted kappa. Absolute agreement is a measure of how often scores from two test replications agree, or how often raters agree on scores given to individual children. Cohen’s kappa quantifies the agreement which exceeds that caused by chance, such that a value of zero would indicate agreement no better than by chance, and a value of one would indicate perfect agreement. In this study, assessing agreement on a five category ordinal scale, Cohen’s quadratic weighted kappa was used for estimation of intra-rater reliability for test replication one and two of each rater (A, B, C, and

D), and for estimation of inter-rater reliability for all six pairs among the four raters (AB, AC, AD, BC, BD, and CD).

Furthermore, the absolute agreement within and between the raters for the BFMF levels, including the sublevels (a) and (b), was calculated.

To interpret the chance-corrected agreement, the following guidelines were used: a value of <0.2 is considered as poor agreement, 0.21–0.4 as fair, 0.41–0.6 as moderate, 0.61–0.8 as good and >0.80 as very good agreement (Altmann, 1991).

Cohen's weighted kappa was calculated using StatXact 11, and the mixed model analyses were carried out using Stata 13. Ninety-five percent confidence intervals (CI) are reported where relevant.

RESULTS

All children (n=79) were classified from the video-recordings by the four raters at both test replications, resulting in a complete data set with no missing values.

In the mixed effect model, the mean score at test session two was 0.009 (p=0.78) lower than at test session one. Hence, there was no clinically important or statistically significant shift between the replications. Furthermore, the overall ICC, reflecting the between rater, within individual, estimate is

$$ICC = \frac{1.496}{1.496 + 0.0018 + 0.243} = \frac{1.496}{1.731} = 0.859$$

In this equation the total variance is the sum of three variance components: variance between children (1.223^2=1.496), variance between raters (0.0427^2=0.0018), and residual variance (0.493^2=0.243). Among the three variance components, the variance between the children is the largest. The variance between raters (equal to 0.0018) is the smallest, and not statistically significant (conservative likelihood ratio test p=0.16).

Measures of intra-rater agreement are shown in Table 2, including classification levels given at test replication one and two. The absolute agreement varied from 0.75 to 0.94, average 0.87. Cohen's weighted kappa varied from 0.91 to 0.97, average 0.95.

The inter-rater reliability for all six pairs of the four raters is shown in Table 3, together with measures of inter-rater agreement. The absolute agreement varied from 0.63 to 0.76, average 0.70. Cohen's weighted kappa varied from 0.86 to 0.93, average 0.90.

The sublevels (a) and (b) were in most cases consistently used, indicated by absolute agreement values varying from 0.71 to 0.91 (average 0.84) for intra-rater reliability and from 0.63 to 0.72 (average 0.60) for inter-rater reliability (Table S1 and S2, online supporting information).

DISCUSSION

The revised BFMF 2 was found to have excellent reliability, as a classification of fine motor capacity, in the included sample of children with CP. The classification levels could be determined from short video recordings (< 5 minutes) of three fine motor tasks regardless of fine motor ability, making the BFMF 2 a feasible classification of fine motor capacity. Furthermore, the figures and the precise descriptions of the fine motor function levels facilitated the use of the classification system. Thus, our results indicate that the BFMF 2 may be a useful instrument for research, as well as for clinical use, giving easily accessible information regarding a child's fine motor function.

It may be considered a strength of the present study that the Consensus-based Standards for selection of Health Measurement Instruments (COSMIN) (Mokkink et al., 2010) was followed. Furthermore, the use of video assessments ensured that the variability of the scoring was unrelated to the child's performance or the instructions. These video recordings were randomly ordered, and the classifications were done individually by each rater at least two weeks apart. It is therefore unlikely that the high reliability was caused by recall bias.

The reliability estimates in this study indicate very good agreement between the raters and must be regarded as very good by any standard. As expected, the ICC estimate of 0.86 is in the same size of order as the average Cohen's weighted kappa of 0.90 (Streiner & Norman, 2015). Furthermore, all the kappa values are clearly higher than the corresponding absolute agreement. This is due to the fact that Cohen's weighted kappa takes into account the magnitude of the disagreement, while the absolute agreement measure counts disagreement equally heavily, independent of the magnitude of the disagreement (Streiner & Norman, 2015). The following observation illustrates this property of Cohen's weighted kappa: Rater A has the highest absolute agreement, with disagreement only on five subjects. But rater D has highest Cohen's weighted kappa. She has disagreement on seven subjects, but with only one classification level difference on these subjects. Since rater A has a difference of two classification levels on two of the subjects, the latter influences more on Cohen's weighted kappa.

High reliability of the BFMF has also been demonstrated in another recently published study from Australia (Randall et al., 2013). In this study, near perfect inter-rater reliability was found with a weighted kappa value of 0.98. Notably, only 20 children were included in this study, which is a low number according to the COSMIN standards for the assessment of the quality of methodological studies (Mokkink et al., 2010). However, this result interpreted together with the comparable findings regarding inter-rater reliability in our study, strongly suggests that BFMF have high inter-rater reliability.

The overall agreement obtained in this study (ICC=0.86), is comparative to what has been reported for the MACS (ICC values varying between 0.75 to 0.98) (Jeevanantham et al., 2015). It has been suggested that an ICC exceeding 0.7 can be regarded as reliable for population-based research, and an ICC exceeding 0.9 is regarded as reliable when it is used clinically (Streiner & Norman, 2015). The reason for the lower demands in population-based studies is that a large sample size will serve to reduce the error of measurement in comparison to group differences. This implies that the BFMF 2 can be used with high confidence in population based studies, comparing results between various raters. For clinical use in individual children, some caution may be needed if the BFMF is classified by different raters, whereas assessments at different time points by the same clinician will be more reliable.

The usefulness of a classification depends on whether or not the different levels are meaningful and understandable, and have clear descriptions (Eliasson et al., 2006). Two recent studies have indicated that the descriptions of the BFMF needed to be clarified (Elvrum et al., 2014; Randall et al., 2013). The original intention of the BFMF was to classify fine motor capacity rather than performance (Elvrum et al., 2014), but this was not explicitly stated in the original description (Beckung & Hagberg, 2002). In the revised BFMF 2 this is clearly stated, securing more similar classifications. Furthermore, the BFMF has been used in modified versions, whereby the child's need for help and adaptations has been added for BFMF levels III-V. This may have made the original intention of the BFMF less clear as a classification of fine motor capacity. Thus, these adaptations were not included in the revised BFMF 2. In the BFMF 2, the child should be given the assistance that is necessary to show his or her best capacity, but the need for assistance should not be classified. Rather, what the child can or cannot do under optimal conditions is classified.

Classification of hand function in children with CP may serve several purposes: (1) to describe common characteristics of hand function in a consistent and meaningful way, (2) to enhance the communication between professionals and parents, and (3) to assist in understanding and identifying the need for intervention and support for children who perform at different levels (Krumlinde-Sundholm, 2008; Rosenbaum et al., 2002). Although the

MACS is widely used as such a measure of manual performance (Jeevanantham et al., 2015), the BFMF may be a useful complement, as it has now been found to be a valid (Elvrum et al., 2014) and reliable instrument describing the fine motor capacity of the two hands separately. In addition, our results indicate that the sub-levels (a) and (b) of the BFMF 2 were quite consistently used, and may provide information about potential differences between the hands. Thereby, the hand function, and specific need for intervention and support can be described in more detail.

LIMITATIONS

Although all CP types were included in this study, the distribution differs somewhat from that of population-based studies (Andersen et al., 2008; Himmelmann et al., 2010), in that the percentage of unilateral spastic CP is lower than expected. Moreover, fewer children than expected were at MACS levels IV-V (Arner et al., 2008) and GMFCS levels IV (Andersen et al., 2008; Carnahan et al., 2007). These limitations notwithstanding, the number of participants is sufficiently large according to COSMIN guidelines (Mokkink et al., 2010), and all CP types and functional levels are represented. Another possible limitation concerning generalizability of our results was that the children were classified mainly by professionals involved in the development of the revised BFMF 2. This degree of familiarization with the BFMF may not be present in the clinic. To account for this, one rater not familiar with the BFMF was included. After brief introduction and practice, she classified all the children with high intra-rater and inter-rater reliability, indicating the usefulness of BFMF 2 also by raters less familiar with the classification. However, others who learn and apply the BFMF 2 according to the manual may be less consistent in their application of the BFMF. To ensure consistency in the classifications regardless of previous experience, the development of an instructional film with examples of children and the rationale for selecting different BFMF levels would be helpful. In this study, only the reliability derived from observations of fine motor capacity was investigated. The BFMF can also be classified from discussions with parents alone, or from a combined use of observations and discussions with parents. Furthermore, for research purposes the BFMF is sometimes classified according to descriptions of hand function in medical and habilitation records (Andersen et al., 2008). The reliability of the BFMF 2 using these approaches needs to be further investigated.

CONCLUSION

The revised BFMF 2 was found to have excellent intra- and inter-reliability. Thus, the use of the BFMF 2 may provide valuable information about the capacity to use each hand separately for holding, grasping and manipulating objects in children with CP. The classification levels could be determined from short video recordings of three fine motor tasks, using the explanatory figures and text added in the revised BFMF 2. This indicates that the BFMF 2 may be a feasible instrument for research as well as for clinically use giving easily accessible information regarding a child's fine motor function.

Declaration of Interest: The authors report no conflict of interest. The authors alone are responsible for the content and writing of this article.

REFERENCES

Altmann, D. (1991). Practical statistics for medical research (Vol. 1st ed.). London: Chapman & Hall/CRC Press.

Andersen, G. L., Irgens, L. M., Haagaas, I., Skranes, J. S., Meberg, A. E., & Vik, T. (2008). Cerebral palsy in Norway: prevalence, subtypes and severity. *Eur J of Paediatr Neurol, 12*(1), 4-13.

Arner, M., Eliasson, A. C., Nicklasson, S., Sommerstein, K., & Hagglund, G. (2008). Hand function in cerebral palsy. Report of 367 children in a population-based longitudinal health care program. *J Hand Surg Am, 33*(8), 1337-1347.

Beckung, E., & Hagberg, G. (2002). Neuroimpairments, activity limitations, and participation restrictions in children with cerebral palsy. *Dev Med Child Neurol, 44*(5), 309-316.

Carnahan, K. D., Arner, M., & Hagglund, G. (2007). Association between gross motor function (GMFCS) and manual ability (MACS) in children with cerebral palsy. A population-based study of 359 children. *BMC Musculoskelet Disord, 8*, 50.

Colver, A. (2006). Study protocol: SPARCLE - a multi-centre European study of the relationship of environment to participation and quality of life in children with cerebral palsy. *BMC Public Health, 6*, 105.

Eliasson, A. C., Krumlinde-Sundholm, L., Rosblad, B., Beckung, E., Arner, M., Ohrvall, A. M., et al. (2006). The Manual Ability Classification System (MACS) for children with cerebral palsy: scale development and evidence of validity and reliability. *Dev Med Child Neurol, 48*(7), 549-554.

Elvrum, A. K., Andersen, G. L., Himmelmann, K., Beckung, E., Ohrvall, A. M., Lydersen, S., et al. (2014). Bimanual Fine Motor Function (BFMF) Classification in Children with Cerebral Palsy: Aspects of Construct and Content Validity. *Phys Occup Ther Pediatr*.

Himmelmann, K., Beckung, E., Hagberg, G., & Uvebrant, P. (2006). Gross and fine motor function and accompanying impairments in cerebral palsy. *Dev Med Child Neurol, 48*(6), 417-423.

Himmelmann, K., Hagberg, G., & Uvebrant, P. (2010). The changing panorama of cerebral palsy in Sweden. X. Prevalence and origin in the birth-year period 1999-2002. *Acta Paediatrica, 99*(9), 1337-1343.

Jeevanantham, D., Dyszuk, E., & Bartlett, D. (2015). The Manual Ability Classification System: A Scoping Review. *Pediatr Phys Ther, 27*(3), 236-241.

Krumlinde-Sundholm, L. (2008). Choosing and using assessments of hand function. In A. C. Eliasson, Burtner, P.A. ed., Improving hand function in children with cerebral palsy: theory, evidence and intervention. London: Mac Keith Press, Vol. 178, 176-197.

Krumlinde-Sundholm, L., Holmefur, M., Kottorp, A., & Eliasson, A. C. (2007). The Assisting Hand Assessment: current evidence of validity, reliability, and responsiveness to change. *Dev Med & Child Neurol, 49*(4), 259-264.

Lien, E., Andersen, G. L., Bao, Y., Gordish-Dressman, H., Skranes, J. S., Vik, T., et al. (2013). Apolipoprotein E polymorphisms and severity of cerebral palsy: a cross-sectional study in 255 children in Norway. *Dev Med Child Neurol, 55*(4), 372-377.

Mokkink, L. B., Terwee, C. B., Patrick, D. L., Alonso, J., Stratford, P. W., Knol, D. L., et al. (2010). The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: an international Delphi study. *Quality of Life Research, 19*(4), 539-549.

Nystrand, M., Beckung, E., Dickinson, H., & Colver, A. (2014). Stability of motor function and associated impairments between childhood and adolescence in young people with cerebral palsy in Europe. *Dev Med Child Neurol, 56*(9), 833-838.

Palisano R, R. P., Bartlett D, Livingston M. (2007). Gross Motor Function Classification System Expanded and Revised. 2011

Randall, M., Harvey, A., Imms, C., Reid, S., Lee, K. J., & Reddihough, D. (2013). Reliable classification of functional profiles and movement disorders of children with cerebral palsy. *Phys Occup Ther Pediatr, 33*(3), 342-352.

Rosenbaum, P. L., Walter, S. D., Hanna, S. E., Palisano, R. J., Russell, D. J., Raina, P., et al. (2002). Prognosis for gross motor function in cerebral palsy: creation of motor development curves. *JAMA, 288*(11), 1357-1363.

SCPE. (2001). Surveillance of cerebral palsy in Europe: a collaboration of cerebral palsy surveys and registers. *Developmental Medicine and Child Neurology, 43*(8), 574-575.

Streiner, D. L., & Norman, G. R. (2015). Health measurement scales : a practical guide to their development and use (5th ed. ed.). Oxford: Oxford University Press.

Öhrvall, A. M., Eliasson, A. C., Lowing, K., Odman, P., & Krumlinde-Sundholm, L. (2010). Self-care and mobility skills in children with cerebral palsy, related to their manual ability and gross motor function classifications. *Dev Med Child Neurol, 52*(11), 1048-1055.

TABLE 1. The distribution of Manual Ability Classification System (MACS) levels and Gross Motor Function Classification System (GMFCS) levels by cerebral palsy (CP) subtype (n=79).

	USCP n=19	CP subtype BSCP n=42	Dyskinetic n=15	Ataxic n=3	Total (%)
MACS levels					
I	6	18	0	0	24 (30)
II	11	11	2	3	27 (34)
III	2	11	3	0	16 (20)
IV	0	1	4	0	5 (6)
V	0	1	6	0	7 (9)
GMFCS levels					
I	15	12	0	1	28 (35)
II	4	20	2	2	28 (35)
III	0	6	2	0	8 (10)
IV	0	4	3	0	7 (9)
V	0	0	8	0	8 (10)

USCP=unilateral cerebral palsy; BSCP=bilateral cerebral palsy; n=number

TABLE 2. The distribution of the five Bimanual Fine Motor Function 2 (BFMF 2) levels assessed by each of the four raters (A, B, C, D) at test replication 1 and 2, and the corresponding intra-rater reliability assessed as absolute agreement in percent, and with the Cohen's quadratic weighted kappa estimates with 95% confidence intervals (CI).

BFMF 2 classification levels								Absolute agreement	Cohen's weighted kappa	
		Test 2							Estimate	(95% CI)
Rater	Test 1	I	II	III	IV	V	Total			
A	I	25	0	0	0	0	25			
	II	0	24	1	1	0	26			
	III	1	0	11	1	0	13			
	IV	0	0	0	9	1	10			
	V	0	0	0	0	5	5			
	Total	26	24	12	11	6	79	0.94	0.96	(0.91 to 1.00)
B	I	25	3	1	0	0	29			
	II	2	16	2	0	0	20			
	III	0	1	6	1	0	8			
	IV	0	0	0	17	0	17			
	V	0	0	0	0	5	5			
	Total	27	20	9	18	5	79	0.87	0.95	(0.92 to 0.99)
C	I	22	1	1	0	0	24			
	II	4	16	2	0	0	22			
	III	0	2	7	2	0	11			
	IV	0	1	4	9	0	14			
	V	0	0	0	3	5	8			
	Total	26	20	14	14	5	79	0.75	0.91	(0.85 to 0.96)
D	I	21	2	0	0	0	23			
	II	1	26	0	0	0	27			
	III	0	3	8	0	0	11			
	IV	0	0	0	10	0	10			
	V	0	0	0	1	7	8			
	Total	22	31	8	11	7	79	0.91	0.97	(0.95 to 0.99)

TABLE 3. The distribution of the five Bimanual Fine Motor Function 2 (BFMF 2) levels assessed by each of the four raters (A, B, C, D) at test replication 1, and the corresponding inter-rater reliability for all six pairs among the four raters (AB, AC, AD, BC, BD, and CD) assessed as absolute agreement in percent, and with the Cohen's quadratic weighted kappa estimates with 95% confidence intervals (CI).

	BFMF 2 classification levels							Absolute agreement	Cohen's weighted Kappa	
Rater		I	II	III	IV	V	Total		Estimate	(95% CI)
				Rater B						
A	I	22	3	0	0	0	25			
	II	7	16	2	1	0	26			
	III	0	1	5	7	0	13			
	IV	0	0	1	8	1	10			
	V	0	0	0	1	4	5			
	Total	29	20	8	17	5	79	0.70	0.90	(0.85 to 0.94)
				Rater C						
A	I	17	7	1	0	0	25			
	II	7	14	3	2	0	26			
	III	0	1	7	5	0	13			
	IV	0	0	0	7	3	10			
	V	0	0	0	0	5	5			
	Total	24	22	11	14	8	79	0.63	0.86	(0.79 to 0.92)
				Rater D						
A	I	19	6	0	0	0	25			
	II	4	18	3	1	0	26			
	III	0	3	6	4	0	13			
	IV	0	0	2	5	3	10			
	V	0	0	0	0	5	5			
	Total	23	27	11	10	8	79	0.67	0.88	(0.83 to 0.94)
				Rater C						
B	I	22	5	2	0	0	29			
	II	2	17	1	0	0	20			
	III	0	0	5	3	0	8			
	IV	0	0	3	11	30	17			
	V	0	0	0	0	5	5			
	Total	24	22	11	14	8	79	0.76	0.91	(0.86 to 0.96)
				Rater D						
B	I	21	8	0	0	0	29			
	II	2	17	1	0	0	20			
	III	0	2	6	0	0	8			
	IV	0	0	4	10	3	17			
	V	0	0	0	0	5	5			
	Total	23	27	11	10	8	79	0.75	0.93	(0.89 to 0.96)
				Rater D						
C	I	18	6	0	0	0	24			
	II	4	17	1	0	0	22			
	III	1	3	6	1	0	11			
	IV	0	1	4	8	1	14			
	V	0	0	0	1	7	8			
	Total	23	27	11	10	8	79	0.71	0.90	(0.84 to 0.95)

Bimanual Fine Motor Function (BFMF) 2nd edition

BFMF levels*

Level I

One hand: manipulates without restrictions. The other hand: manipulates without restrictions or limitations in more advanced fine motor skills

Level II

(a) One hand: manipulates without restrictions. The other hand: only ability to grasp or hold
(b) Both hands: limitations in more advanced fine motor skills

Level III

(a) One hand: manipulates without restrictions. The other hand no functional ability
(b) One hand: limitations in more advanced fine motor skills. The other hand: only ability to grasp or worse

Level IV

(a) Both hands: only ability to grasp
(b) One hand: only ability to grasp. The other hand: only ability to hold or worse

Level V

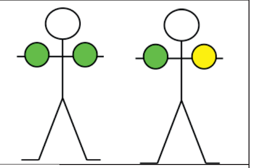
Both hands: only ability to hold or worse

*The same description of the BFMF levels as in the original version by Beckung and Hagberg (2002).

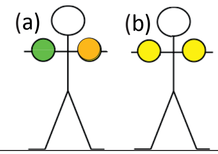
Explanatory figures with color-coded text:

- Without restriction (green)
  - In-hand-manipulation: with speed and precision
  - Grasp all kind of objects with speed and precision
- Restriction in advanced fine motor skills (yellow)
  - In-hand manipulation: reduced speed and precision
  - Grasps objects from table; reduced speed and precision.
- Can grasp and hold, no in-hand manipulation (orange)
  - No in-hand manipulation, may manipulate against table or body
  - Grasps selected objects from the table and other objects from an adapted position, reduced speed and precision
- May hold (red)
  - No manipulation of objects
  - Cannot grasp objects from the table, may grasp a few objects from an adapted position
  - May hold object placed in hand

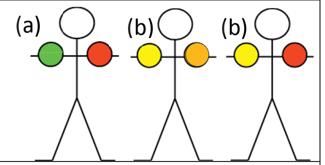
BFMF I



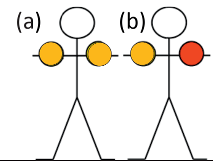
BFMF II



BFMF III



BFMF IV



BFMF V

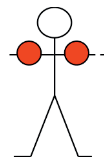


Table S1: Classification of Bimanual Fine Motor Function 2 (BFMF 2) levels and sublevels (a) and (b) at test session 1 and 2, for each of the four raters (A, B, C, D). The grey areas show the agreement between the sublevels (a) and (b) for classification levels II-IV. In addition the absolute agreement for all classification levels and sublevels are reported.

		BFMF 2 classification levels								Absolute agreement
		Test 2								
Rater	Test 1	I	IIa	IIb	IIIa	IIIb	IVa	IVb	V	
A	I	25	0	0	0	0	0	0	0	
	IIa	0	15	0	0	0	1	0	0	
	IIb	0	0	9	0	1	0	0	0	
	IIIa	0	0	0	6	0	0	0	0	
	IIIb	1	0	0	1	4	1	0	0	
	IVa	0	0	0	0	0	5	0	0	
	IVb	0	0	0	0	0	1	3	1	
	V	0	0	0	0	0	0	0	5	0.91
B	I	25	1	2	0	1	0	0	0	
	IIa	1	14	0	1	1	0	0	0	
	IIb	1	0	2	0	0	0	0	0	
	IIIa	0	0	0	4	0	0	0	0	
	IIIb	0	0	1	0	2	1	0	0	
	IVa	0	0	0	0	0	8	0	0	
	IVb	0	0	0	0	0	2	7	0	
	V	0	0	0	0	0	0	0	5	0.85
C	I	22	1	0	0	1	0	0	0	
	IIa	2	11	0	0	1	0	0	0	
	IIb	2	0	5	0	1	0	0	0	
	IIIa	0	0	0	2	2	0	0	0	
	IIIb	0	1	1	0	3	2	0	0	
	IVa	0	0	1	0	3	7	0	0	
	IVb	0	0	0	0	1	1	1	0	
	V	0	0	0	0	0	1	2	5	0.71
D	I	21	0	2	0	0	0	0	0	
	IIa	0	19	0	0	0	0	0	0	
	IIb	1	0	7	0	0	0	0	0	
	IIIa	0	0	0	4	1	0	0	0	
	IIIb	0	1	2	0	3	0	0	0	
	IVa	0	0	0	0	0	8	0	0	
	IVb	0	0	0	0	0	0	2	0	
	V	0	0	0	0	0	0	1	7	0.90

Table S2: Classification of Bimanual Fine Motor Function version 2 (BFMF 2) levels and sublevels (a) and (b) at test session 1 for all six pairs among the four raters (AB, AC, AD, BC, BD, and CD). The grey areas show the agreement between the sublevels (a) and (b) for classification levels II-IV. In addition Cohen's unweighted kappa (kappa) values are estimated only taking the sublevels (a) and (b) into account, with 95% confidence intervals (CI).

		BFMF 2 classification levels								Absolute agreement
		I	IIa	IIb	IIIa	IIIb	IVa	IVb	V	
		Rater B								
Rater A	I	22	2	5	0	0	0	0	0	
	IIa	2	13	1	1	0	0	0	0	
	IIb	1	0	2	0	0	0	0	0	
	IIIa	0	0	0	4	0	0	0	0	
	IIIb	0	0	2	0	1	1	0	0	
	IVa	0	0	0	0	5	3	0	0	
	IVb	0	1	0	1	1	1	4	1	
	V	0	0	0	0	0	0	1	4	0.67
		Rater C								
Rater A	I	17	3	4	0	1	0	0	0	
	IIa	3	10	0	0	2	1	0	0	
	IIb	4	0	4	0	1	1	0	0	
	IIIa	0	1	0	4	0	0	1	0	
	IIIb	0	0	0	0	3	4	0	0	
	IVa	0	0	0	0	0	5	0	0	
	IVb	0	0	0	0	0	0	2	3	
	V	0	0	0	0	0	0	0	5	0.63
		Rater D								
Rater A	I	19	3	3	0	0	0	0	0	
	IIa	0	14	0	0	1	1	0	0	
	IIb	4	0	4	0	2	0	0	0	
	IIIa	0	2	0	4	0	0	0	0	
	IIIb	0	0	1	0	2	4	0	0	
	IVa	0	0	0	1	1	3	0	0	
	IVb	0	0	0	0	0	0	2	3	
	V	0	0	0	0	0	0	0	5	0.67
		Rater C								
Rater B	I	22	1	4	0	2	0	0	0	
	IIa	2	13	1	0	1	0	0	0	
	IIb	0	0	3	0	0	0	0	0	
	IIIa	0	0	0	4	0	0	0	0	
	IIIb	0	0	0	0	1	3	0	0	
	IVa	0	0	0	0	2	6	0	0	
	IVb	0	0	0	0	1	2	3	3	
	V	0	0	0	0	0	0	0	5	0.72
		Rater D								
Rater B	I	21	4	4	0	0	0	0	0	
	IIa	1	14	1	0	1	0	0	0	
	IIb	1	0	2	0	0	0	0	0	
	IIIa	0	1	0	3	0	0	0	0	
	IIIb	0	0	1	0	3	0	0	0	
	IVa	0	0	0	1	1	6	0	0	
	IVb	0	0	0	1	1	2	2	3	
	V	0	0	0	0	0	0	0	5	0.71
		Rater D								
Rater C	I	18	4	2	0	0	0	0	0	
	IIa	1	12	0	0	1	0	0	0	
	IIb	3	0	5	0	0	0	0	0	
	IIIa	0	1	0	3	0	0	0	0	
	IIIb	1	2	0	0	3	1	0	0	
	IVa	0	0	1	1	2	7	0	0	
	IVb	0	0	0	1	0	0	1	1	
	V	0	0	0	0	0	0	1	7	0.71

Paper III

Outcome measures evaluating hand function in children with bilateral cerebral palsy: A systematic review

ANN-KRISTIN G. ELVRUM,[1,2] RANNEI SÆTHER,[2,3] INGRID I. RIPHAGEN,[4] TORSTEIN VIK[2]

[1]Department of Clinical Services, St. Olavs Hospital, Trondheim University Hospital, Norway

[2]Department of Laboratory Medicine, Children's and Women's Health, Faculty of Medicine, Norwegian University of Science and Technology, Trondheim, Norway

[3]Department of Paediatrics, St. Olavs Hospital, Trondheim University Hospital, Norway

[4]Department of Cancer Research and Molecular Medicine, Norwegian University of Science and Technology, Trondheim, Norway

Correspondence to
Ann-Kristin G. Elvrum,
Department of Laboratory Medicine, Children's and Women's Health, Faculty of Medicine,
Norwegian University of Science and Technology, Post-box 8905,
NO-7491 Trondheim, Norway

E-mail: ann-kristin.elvrum@ntnu.no

Word count: 5264

Abstract

Aim: To review outcome measures used to evaluate hand function, with emphasis on manual capacity and performance, in children with bilateral cerebral palsy (CP), to describe the content and measurement properties of such measures, and to investigate the quality of the studies that have examined these properties.
Methods: Embase, MEDLINE, PubMed and CINAHL were searched. The COSMIN-criteria (Consensus-based standards for selection of health measurement instruments) were used to assess the quality of studies and the Terwee criteria were used to assess the result of the studies.
Results: Five hand function measures were identified from 16 papers. The strongest level of evidence for aspects of validity and reliability was found for the Melbourne Assessment 2, assessing unimanual capacity, and for the questionnaire ABILHAND-Kids, assessing perceived manual ability in daily activities. However, evidence for the responsiveness of these measures is missing.
Interpretation: Further high quality studies providing evidence for responsiveness, as well as for additional aspects of validity and reliability of the Melbourne Assessment 2 and the ABILHAND-Kids, are needed. Furthermore, there is a need to develop appropriate outcome measures evaluating how children with bilateral CP use their hands when handling objects in bimanual tasks.

What this paper adds:

- Only five measures assessing hand function in children with bilateral CP were identified.
- The strongest level of evidence was found for Melbourne Assessment 2 and ABILHAND-Kids.
- Further research investigating responsiveness of these measures is required.
- No measure was identified assessing how the hands are used in bimanual tasks.

Running foot: Hand function measures in bilateral CP

Outcome measures with sound measurement properties are crucial for appropriate planning and evaluation of interventions as well as for the assessment of the development of children with chronic health conditions, such as cerebral palsy (CP). The assessment of a child's development can be used to counsel families regarding prognostic factors, as well as informing about the child's strengths and limitations in the attribute measured and expected treatment outcomes. In addition changes over time associated with service provision can be documented allowing enhanced clinical decision making and appropriate allocation of resources.[1]

The selection of appropriate outcome measures can be facilitated by systematic appraisal of the content and measurement properties of the instruments in question.[2] An outcome measure used for evaluation should address the domain of concern, be valid and reliable in the population of interest, be responsive to change, and have good clinical utility.[3]

Recent reviews have confirmed that valid and reliable outcome measures exist to evaluate hand function in children with unilateral cerebral palsy[4-8] and a few systematic reviews have summarized evidence on parent reported measures of functional hand use[9] and assessment tools and classification systems of the upper extremity in children with CP in general.[10,11] For children with bilateral CP, however, there is little evidence on efficacy of interventions targeting their upper extremity function,[12] and this may in part be due to lack of appropriate outcome measures. This is of concern since more than 60% of children with bilateral CP have decreased hand function.[13] Thus, it is of upmost importance to identify and appraise outcome measures evaluating hand function in these children. According to the International Classification of Functioning, Disability and Health (ICF) conceptual framework [14] hand function can be assessed according to the child's functional *performance* (what the child usually does) or by testing what a child is able to do on request (*capacity*). What a child *can do* and *does do* may not always be equivalent. Thus, outcome measures assessing both these concepts are needed for treatment planning and evaluation.[15]

The overall aim of this systematic review was therefore to identify available outcome measures evaluating hand function in children with bilateral CP, with emphasis on manual capacity and performance, and to appraise the current level of evidence for the measurement properties of these instruments.

Methods

A protocol describing the current review has been recorded and published on the PROSPERO register database, registration number: CRD42015019544.

Search strategy

Searches were carried out in Embase (through OvidSP edition 1980 and onwards), MEDLINE (through OvidSP edition 1946 and onwards), PubMed and CINAHL (through Ebscohost) until 10 June 2015. The first search aimed to identify hand function measures in children with bilateral CP. The search strategy used text words and medical subject headings (MeSH, EMtree). Queries consisted of Boolean combinations of search term groups representing the construct of interest (hand function), the target population (children with diplegic, quadriplegic, dyskinetic and ataxic CP), instruments and measurement properties.

Subsequently, the names of the outcome measures identified through the first search were used in a complementary search, which aimed to identify additional studies of the measurement properties of the hand function measures. Boolean combinations of search term groups representing 'instrument name in title' and 'validation' were applied. Details of the searches in the databases are presented in Appendix S1 (online supporting information).

Furthermore, references in the included articles and reviews were checked. Papers assessing the measurement properties of the hand function instruments were included for quality analysis.

Study selection

Titles and abstracts were initially screened for relevance by the first author (AKGE), with the third author (IR) screening excluded titles and abstracts to ensure that no relevant papers were omitted. Secondly full text papers were reviewed by the first (AKGE) and second author (RS) according to the inclusion and exclusion criteria. Cases of disagreement were discussed until consensus was reached. To be included, studies had to meet the following criteria: (1) the study investigated measurement properties of outcome measures where the primary intention was to evaluate hand function according to how children handle objects, i.e. within the activity component of the ICF.[16] Studies of outcome measures evaluating a combination of concepts within the activity and body function components of the ICF were also included; (2) the study population consisted of children with bilateral CP of the spastic, dyskinetic or ataxic types in the age range 0-18 years; (3) the measure was designated for use in a clinical setting, without the need for laboratory equipment; and (4) the paper describing the study was a full-text original article (e.g. not a review or manual). Furthermore, studies were excluded if they were: (1) not published in English; (2) evaluated measurement properties of classification systems or of measures primarily assessing participation, body function, body structures, school-functioning, self-care or health-related quality of life; or (3) children with bilateral CP comprised less than 30% of the total population.

Data extraction and quality assessment

The data extraction and quality assessment consisted of several steps. First, the general characteristics of the outcome measures, such as focus, scale construction, standardization and clinical utility, were extracted by the first author (AKGE) in a review table, adapted from the CanChild Outcome Measures Rating Form.[17] To identify the focus of each measure the items were linked to relevant categories in the body function and structure, activity and participation domains of the ICF framework,[18] and it was considered whether capacity (what the child “can do”) or actual performance (what the child “does do”) was assessed. Scale construction and standardization of each measure was reviewed according to item selection and to whether the measure was norm-referenced or criterion-referenced. In a norm-referenced test the child’s performance is compared to the average performance of a normative sample of typically developed age-matched peers. These tests are usually used for diagnostic purposes, since they provide information about how much the child’s performance deviates from the mean score of a normative sample. In a criterion-referenced test the child’s performance is measured according to clearly stated criteria describing particular levels of performance. These tests are generally more useful for establishing treatment goals and measuring change, since they can report progress related to specific criteria instead of reporting deviations from an age-related norm.[19,20] Furthermore, considerations of clinical utility of the measures concerned time for administration and required examiner qualifications.

Second, descriptive characteristics of the population and the measurement properties reported in the clinimetric studies were extracted. Third, the methodological quality of studies of measurement properties of the included outcome measures was evaluated independently by the first (AKGE) and second author (RS) according to the COnsensus-based Standards for the selection of health Measurement INstruments (COSMIN).[2] In case of disagreement, the results were discussed until consensus was reached.

The COSMIN checklist consists of the following measurement properties: internal consistency, reliability, measurement error, content validity, construct validity (subdivided into structural validity, hypotheses testing, and cross-cultural validity), criterion validity and responsiveness. In addition, general requirements for studies that applied Item Response Theory (IRT) models were appraised when applicable.[2] Each measurement property was evaluated separately according to between 5 to 18 items rating the design requirements and statistical methods according to a four-point rating scale (i.e. poor, fair, good, and excellent). Subsequently, an overall score for the methodological quality per measurement property in each study was determined by the lowest rating of any of the items ('worst score counts' method).[21]

Some of the items in the COSMIN checklist require subjective judgment[21] and in this review two adaptations clarifying the quality scorings in COSMIN were made. The first adaptation was related to the reporting of missing items. Since the majority of the included studies either did not indicate the number of missing items and/or did not indicate how missing items were handled, we decided to exclude these items from the quality ratings to be able to detect additional strengths and/or limitations in the quality of the included studies. The second adaptation was related to sample size. In the COSMIN checklist, studies with samples of less than 30 participants are rated as "poor". However, in a systematic review many small studies together can provide enough evidence. Thus, in line with previous systematic reviews[22,23] and after consultation with authors of the COSMIN checklist, we decided to omit the sample size item from the quality assessment and instead account for sample size at the best evidence synthesis stage described below.

Based on quality criteria proposed by Terwee et al[24] the first author (AKGE) rated the results of the measurement properties for each study as positive (+), indeterminate (?), or negative (-) (Table S1, online supporting information). In addition the level of overall evidence for each outcome measure was estimated by taking into consideration a) the number of studies, b) the methodological quality of the studies and c) the consistency of their results in a similar manner to that recommended by the Cochrane Back Review Group.[25] Only studies of "fair", "good" or "excellent" methodological quality were included in the best evidence synthesis.[23,26] The possible levels of evidence are "strong", "moderate", "limited", "conflicting", or "unknown" (Table S2, online supporting information). To account for sample size, the level of evidence was rated as "strong" when the total sample size of included studies was ≥100, "moderate" for a total sample size between 50 and 99, "limited" for a total sample size between 30 and 49, and "unknown" when sample size was <30.[21-23]

Results

Study selection

The selection process used to identify relevant papers is shown in the flowchart (See Fig. 1). A total of 934 unique records of possible interest were revealed in the database searches. After screening of titles and abstracts of identified papers, 212 records were included for a full text check. After a full text check, a further 197 papers were excluded. Finally, one paper was added after screening the references, reviews and published abstracts from conferences, and a total of 16 papers and five hand function measures were included for quality assessment. The included hand function measures were one parent-reported questionnaire; the ABILHAND-Kids,[27] and four standardized tests; the Erhardt Developmental Prehension Assessment (EDPA),[28] the Peabody Developmental Motor Scales (PDMS),[29] the Quality of Upper Extremity Skills Test (QUEST),[30] and the Melbourne Assessment of Unilateral Upper Limb Function (MA).[31] For the MA and the PDMS revised versions of the measures have been published, the Melbourne Assessment 2 (MA2)[32] and the Peabody Developmental Motor Scales Second Edition (PDMS-2).[33] These versions are slightly different from the original versions and are therefore treated separately in the tables and in the quality appraisal of the current level of evidence for the measurement properties. The PDMS/PDMS-2 consists of a Gross Motor Scale and a Fine Motor Scale (FMS), but in the current review only measurement properties of the Fine Motor Scale are reported.

In line with the inclusion criteria, some frequently used hand function measures were excluded from our review (Table S3, online supporting information). For example, two timed tests measuring unilateral capacity, the Box and Blocks test of manual dexterity (B&B)[34,35] and the Jebsen-Taylor Test (JTT)[36] were excluded since measurement properties of these measures have not been tested in individuals with bilateral CP. Furthermore, two commonly used norm-referenced tests of fine motor function, the Bayley Scales of Infant and Toddler Development (Bayley)[37,38] and the Bruininks–

Oseretsky Test of Motor Proficiency (BOTMP),[39] were excluded for the same reason. Other measures were excluded because their primary purpose was not to assess hand function; rather the hand function assessment was part of a wider assessment of self-care, school-functioning, participation or quality of life. Among these measures were the Children's Assessment of Participation and Enjoyment, the Paediatric Evaluation of Disability Inventory, the Vineland adaptive behaviour scales, fine motor 2nd edition and the Pediatric Outcomes Data Collection Instrument. The measurement properties of the new outcome measure Children's Assessment of Participation with hands has so far only been investigated in a study in which individuals with CP comprised only 7% of the study population and the CP subtypes were not specified.[40]

Characteristics of the hand function measures

The five included hand function measures varied in focus, attribute measured, scale construction and standardization, as well as clinical utility (see Table I). The focus of all items included in the questionnaire (ABILHAND-Kids) was within the activity domain of the ICF, while the four standardized tests included items that were within both the activity and the body functions components. Furthermore, the ABILHAND-Kids measures manual ability according to ease or difficulty performing activities requiring the use of one or both hands in daily life activities (i.e. perceived performance), while the four standardized tests measures the child's best ability to use the hands (i.e. capacity). The attribute measured, scale construction and standardization of the included outcome measures are presented below in chronological order according to publishing date.

The EDPA was published in 1982[28] with a 2nd edition published in 1994,[41] but the test criteria were unchanged. The test aims to measure involuntary and voluntary movement patterns from birth through 15 months, and pencil grasp and drawing skills from one to six years of age, in infants and children with CP or other motor disorders. The test criteria for grading the child's abilities were based on developmental norms extracted from evaluation scales and child developmental literature, but these norms have not been tested empirically. The primary purpose of the EDPA was to enable evaluation of neuro-developmental interventions. Thus, the selection of test items was based on neurodevelopmental theories and the test includes assessment of reflexive movement patterns, as well as voluntary reach, grasp, manipulation and release, in each hand separately.

The PDMS is a norm-referenced, standardized test that was published in 1983[29] with the release of a second version (PDMS-2) in 2000.[33] It aims to measure motor abilities that develop early in life (PDMS: birth through 83 months, PDMS-2: birth through 71 months) relative to age-related typical development, and thereby identify children with delayed motor development. Thus, it is first and foremost a diagnostic/discriminative measure. The test includes a separate fine motor scale that is divided into four subscales for the PDMS (grasping, hand use, eye-hand coordination and manual dexterity) and two subscales for the PDMS-2 (grasping and visual motor integration). Both unimanual and bimanual tasks are included. Unimanual tasks are scored based on how they are performed with the preferred hand, while bimanual tasks are scored according to task accomplishment (i.e. ability to fold a paper, ability to cut a circle or a square).

The QUEST is a criterion-referenced test that was published in 1992[30] to enable evaluation of interventions targeting movement patterns in the upper limbs. It aims to measure unilateral movement quality and grasping ability in children with CP aged 18 months to 8 years, and measurement properties of the QUEST have been investigated for children in MACS levels I-V. The QUEST consists of four domains: (1) dissociated movements, (2) grasp, (3) weight bearing and (4) protective extension. The selection of test items was based on neurodevelopmental theories regarding essential components of hand function, and nearly 80% of the test items assess body functions related to posture, range of motion and weight bearing. Seven items assess the ability to grasp and release with each hand separately. None of the items assess ability to use both hands together.

The MA is a criterion-referenced, standardized test that was first published in 1999[31] with a revised second version (MA2) released in 2012.[32] The test intends to discriminate between different levels of hand function and to evaluate treatment. The MA was designed to measure the quality of unilateral upper limb movements in children between the ages of five and 15 years with CP or other neurological impairments. With the MA2 the age range was expanded to include children from the age of 2.5 through 15 years. Furthermore, Rasch analysis indicated the need to split the test into four discrete subscales: (1) range of motion, (2) accuracy, (3) dexterity, and (4) fluency. The dexterity subscale evaluates ability to grasp, release and manipulate in each hand separately, i.e. within the activity component of the ICF. The other subscales evaluate body functions only (range of motion and fluency subscales), or both (accuracy subscale). All items are scored from observations of unimanual tasks including capacity to reach, grasp, release and manipulate. The test does not provide information about the ability to use both hands together.

The ABILHAND-Kids is a parent reported questionnaire that was developed in 2004 to measure and evaluate manual ability in children with CP aged 6-15 years. [42] Manual ability was defined as performance of daily activities requiring the use of the upper limbs, and the ABILHAND-Kids includes items such as taking off a T-shirt, filling a glass of water, or putting on a backpack. The parent records whether they perceive that each item is impossible, difficult, or easy to complete for their child without human or technical assistance. The item is scored on the completion of the task regardless of how the item is completed. Thus, the ABILHAND-Kids does not provide information regarding whether one or two hands are used, and fine motor components of hand function, such as ability to reach, grasp, manipulate or release, are not considered.

The descriptions of the populations with bilateral CP in which the measures have been applied varied. Only three studies investigating measurement properties of the QUEST reported Manual Ability Classification System (MACS) levels,[43-45] suggesting that the QUEST is appropriate for children in MACS levels I-V. For the other outcome measures studies investigating measurement properties reported the Gross Motor Function Classification System (GMFCS) levels[42,46,47] or clinical descriptions such as the severity of CP (mild, moderate, and severe)[48-51] or CP subtype.[52-55] Several of the studies not reporting on MACS levels were published before this classification system became available. However, according to the content of the measures we assumed that the EDPA, the MA/MA2, and the ABILHAND-Kids would be appropriate for children in MACS levels I-V, while the norm-referenced PDMS/PDMS-2 FMS would probably be most relevant for children in MACS levels I-II.[20]

Clinical utility was considered according to time needed to complete the assessments and examiner qualifications. All the standardized tests could be completed in about 30-45 minutes and training or experience was recommended to ensure reliable administration and scoring for these measures. The parent-reported questionnaire, ABILHAND-Kids, could be completed in 10-15 minutes.

Quality of the studies

The quality appraisal of the measurement properties was performed according to the COSMIN four-point scale and resulting details are summarized in Tables II and III. In the 16 papers 30 reliability properties (internal consistency, intra-/interrater and test-retest, and measurement error) and 12 validity properties (content, construct; structural validity and hypotheses testing) were assessed, while responsiveness was assessed only in one study.[48] One of the included studies described content validity and construct validity of a Modified MA for neurologically impaired children aged two to four years.[50] The Modified MA is not available and the results regarding construct validity of this test version is therefore not included in the current review. However, the described content validity of the Modified MA was incorporated in the development of the MA2 test version,[51] and the study[50] was therefore included in the quality appraisal of the measurement properties of the MA2.

The quality of how the reliability properties had been investigated was rated as 'poor' (n=6), 'fair' (n=8), 'good' (n=5) and excellent (n=11) (See Table II). The main reasons for the low quality scores were lack of factor analysis or important methodological flaws.[44,45,49,52] Two of the three studies investigating internal consistency using the Rasch measurement model was rated as "excellent",[42,51] while the other was rated as "poor" due to extensive use of repeated assessments, as well as few items for the subscale analyses.[44] In the latter study as many as 51 out of 94 children contributed with between two and four assessments performed 4-10 weeks apart, most likely violating independency in the data set.[56]

Furthermore, the quality of how validity properties had been investigated was rated as "poor" (n=3),"fair" (n=2), and "excellent" (n=7) (See Table III). The reasons for low quality scores for construct validity were poor description of comparator instrument,[52] as well as methodological issues such as lack of a priori formulated hypothesis[47] and use of repeated assessments.[44] All studies of content validity were rated as excellent.[42,50,52,54,57] The quality of the only study of responsiveness was rated as "fair".[48]

Results of studies

The ratings of the reported results of the studies, according to Terwee criteria are summarized in Tables II and III. The results of the studies for reliability (n=30) were rated as 'positive' (n=24), 'indeterminate' (n=3), and 'conflicting' (n=3; studies had both negative and positive results), the results of the studies for validity (n=12) were rated as 'positive' (n=10) or 'indeterminate' (n=2), and the study of responsiveness was rated as 'positive'.

Data synthesis

Table IV shows an overall summary of measurement properties of the five included hand function measures, as well as the level of evidence of the properties. Four of the measures had strong evidence for content validity; the EDPA, the QUEST, the MA/MA2 and the ABILHAND-Kids, while the content validity of the PDMS/PDMS-2 FMS has not been studied specifically for children with bilateral CP.

For the EDPA there was no evidence for other measurement properties, besides content validity.[54] Inter-rater reliability has also been reported for the EDPA, but with unknown evidence due to a very limited sample size.[54]

The PDMS-2 FMS was the only measure where responsiveness has been studied, however with limited evidence.[48] In addition aspects of reliability, such as test-retest and measurement error have been studied, but were only found to have moderate or limited evidence.[48]

For the QUEST strong positive evidence was found for content validity,[57] and moderate positive evidence was found for interrater reliability of total scores when several small studies were pooled.[43,57,58] However, for the Grasp subdomain of the QUEST the results were more conflicting. Furthermore, limited evidence was found for intra-rater,[43,58] test-retest[57,58] and hypotheses testing,[57] while the results from the studies investigating internal consistency and structural validity were not included due to poor quality of the studies.[44,45]

The MA2 and the ABILHAND-Kids were found to have strongest evidence, with strong level of evidence for internal consistency and structural validity, [42,51] in addition to content validity.[42,50] Furthermore, moderate evidence was found for test-retest properties of the ABILHAND-Kids.[42] Other reliability and validity properties, as well as responsiveness, have not yet been investigated for these measures in children with bilateral CP.

Criterion validity was not studied for any of the outcome measures.

Discussion

Our systematic review identified five hand function measures, where measurement properties have been evaluated in children with bilateral CP from 0-18 years of age. The MA2, measuring unimanual capacity, and the ABILHAND-Kids, assessing perceived manual ability, had the strongest level of evidence for aspects of both reliability and validity, but no evidence of responsiveness. None of the identified outcome measures evaluated actual bimanual performance (i.e. spontaneous handling of objects requiring the use of both hands) in children with bilateral CP.

One of the strengths of our review was the inclusion of studies of measurement properties for a relatively homogeneous population. The CP diagnosis in general includes a heterogeneous group of disorders of movement and posture causing activity limitations. Thus, outcome measures validated for one subtype (e.g. unilateral spastic CP) may not apply to children within other subtypes. It is therefore essential that measurement properties of relevant outcome measures for the various subtypes are investigated.

Through the combined assessment of the methodological quality and the results of the included studies (i.e. the reported measurement properties of each measure), the current review offer an overview of the overall evidence concerning measurement properties of outcome measures used to evaluate hand function in children with bilateral CP. To be able to trust the results regarding measurement properties of an outcome measure, it is important that the methodological quality of the studies of measurement properties is appropriate.[21] Although the majority of the included studies reported positive results for the investigated measurement properties, the best evidence synthesis indicated limited evidence due to poor to fair methodological quality in nearly half of those studies. The rating of the methodological quality of the included studies according to the COSMIN was in this regard very useful.[21]

There are some potential limitations to this review. Articles were only included if they were published in English and, therefore, some papers investigating measurement properties of identified outcome measures, may have been missed. A further limitation may be the exclusion of studies on measurement properties where the main aim was not to assess measurement properties of the measure in question, for example intervention studies reporting a short note on assessment of reliability of the outcome measures used. This may have caused some loss of information regarding measurement properties for some of the outcome measures. However, it is unlikely that the inclusion of such studies would have changed our findings significantly. Furthermore, we identified only one study[48] assessing responsiveness according to the COSMIN standards, which state that the treatment effect measured by a specific outcome measure and the responsiveness of the same measure should not be based on the effect size in the same study.[59] In line with these standards we did not include studies where the primary aim was to assess the effect of interventions.

The relatively large number of outcome measures identified compared to the few included, may reflect that many different measures are used to assess hand function, but very few of these are actually validated for children with bilateral CP. Among the excluded outcome measures were two timed tests, the JTT and the B&B, assessing unimanual speed and dexterity. Speed is considered to be an important aspect of effective hand use[60] that is not specifically evaluated by the outcome measures that are included in this systematic review. Thus, high quality studies investigating measurement properties of these measures in children with bilateral CP are warranted.

In the following paragraphs the pros and contras of the included outcome measurements will be discussed. For the norm-referenced test PDMS-2 FMS, test-retest reliability, measurement error and responsiveness have been investigated in children with bilateral CP.[48] However, the PDMS-2 FMS has never been validated for this population. Thus, the use of this test for evaluative purposes in children with bilateral CP may be questioned. The PDMS-2 FMS, and other norm-referenced test (i.e. the Bayley and the BOTMP) have been found to be useful in identifying children whose fine motor abilities deviate from the norm, for example for diagnostic or discriminative purposes.[19,20] However, the use of norm-referenced tests as evaluative measures in children with permanent motor disorders, such as CP, may not be appropriate since these children may never catch up with the increasing demands of age norm-referenced test criteria, despite improvement in skills.[20]

Two of the included outcome measures, the EDPA and the QUEST, were developed to enable evaluation of neuro-developmental interventions.[30,54,61] Thus, the selection of test items for both these measures included assessments of posture, as well as reflexive and voluntary movement patterns. The use of neuro-developmental interventions is debated, and a recent review found no evidence to continue the use of this approach.[12] Consequently, the use of outcome measures basing their content on the same theories should also be questioned. For the EDPA we found no evidence of reliability and validity, except for content validity. Thus, continued use of this measure in children with bilateral CP is probably not recommended. For the QUEST we found limited to moderate evidence for test-retest, intra- and inter-rater reliability. However, when the QUEST was investigated by the use of Rasch modelling, the test was not found to have adequate measurement properties.[44] Unfortunately this study could not be included in the best evidence synthesis because of poor quality. Still, we believe that the results from this study can be taken into account, indicating that the QUEST is maybe not an appropriate measure of upper extremity quality of movement in children with bilateral CP, as it is now. This is also indicated in another recent study of reliability of the QUEST.[45]

Consequently, only two of the included outcome measures, the MA2 and the ABILHAND-Kids, can be recommended for use in children with bilateral CP. Both these measures were found to have strong evidence for aspects of both validity and reliability. The criterion-referenced MA2 has a solid foundation to its measurement properties based on refinement of the original MA through Rasch modelling. Use of the Rasch measurement model , both in the development of a new measure and/or in the refinement of existing measures, is becoming a preferred way to construct outcome measures for use in rehabilitation.[62] For the MA, Rasch analysis showed the need to split the original MA into four discrete subscales, leading to the development of the MA2. Thus, the measurement properties of the original MA do not apply to the MA2, and further research is required to determine the discriminative and evaluative abilities of the four MA2 subscales. So far, the sensitivity and the ability of the MA2 to detect small, but clinically important changes, is unknown. The four subscales of the MA2 may highlight relative areas of strength and weakness of movement quality regarding range of motion, accuracy, fluency and dexterity in each upper limb separately. However, the MA2 does not provide information about the ability to use both hands together. Thus, other outcome measures evaluating bimanual performance are required to provide necessary and complementary information regarding hand function in children with bilateral CP.

The ABILHAND-Kids is a parent-reported questionnaire evaluating perceived manual performance in children aged 5-16 years with CP. The questionnaire is indicated when the aim is to evaluate ease or difficulty of bimanual performance in everyday activities, but it cannot be used to define functional goals since no information regarding actual use of the hands is provided (i.e. whether one or both hands are used). Through Rasch-modelling the ABILHAND-Kids has been found to have strong evidence for internal consistency and structural validity. Furthermore, moderate evidence for test-retest reliability has been demonstrated. In addition some studies not intended to focus on measurement properties of ABILHAND-Kids, and thus not included in this systematic review, may suggest that the ABILHAND-Kids is able to discriminate between different levels of hand function.[63,64] However, this warrants further investigation together with research on other aspects of reliability, as well as on its responsiveness to change.

Strikingly, no outcome measure exists to assess how children with bilateral CP actually use the hands together when handling objects. In contrast, for children with unilateral CP the Assisting Hand Assessment (AHA) fulfills this aspect by measuring spontaneous and effective use of the affected hand in bimanual activities.[15] The AHA is a Rasch-build measure that has proven to be valid, reliable and sensitive to change.[65,66] It is therefore used to monitor upper extremity development and to evaluate interventions in children with unilateral CP.[67] Consequently, it would be of great interest if the AHA could be modified for use to evaluate actual bimanual performance in children with BCP.

Conclusion

In this systematic review we identified five measures assessing hand function of children with bilateral CP where aspects of reliability, validity and responsiveness had been studied. Of these, the MA2 measuring unimanual movement quality, and the ABILHAND-Kids, measuring perceived manual performance, had the strongest levels of evidence for aspects of both validity and reliability. However, further research is required to determine the responsiveness of these measures. Furthermore, we found it of particular concern that we were not able to identify outcome measures assessing the bimanual performance in children with bilateral CP. Thus, there is a need to develop outcome measures that validly and reliable assess how children with bilateral CP use the hands together when handling daily life objects.

References

1. Majnemer A, Mazer B. New directions in the outcome evaluation of children with cerebral palsy. *Semin Pediatr Neurol*. 2004; **11**: 11-7.
2. Mokkink LB, Terwee CB, Knol DL, et al. The COSMIN checklist for evaluating the methodological quality of studies on measurement properties: a clarification of its content. *BMC Medical Research Methodology*. 2010; **10**: 22.
3. Mokkink LB, Terwee CB, Knol DL, et al. Protocol of the COSMIN study: COnsensus-based Standards for the selection of health Measurement INstruments. *BMC Med Res Methodol*. 2006; **6**: 2.
4. Gilmore R, Sakzewski L, Boyd R. Upper limb activity measures for 5- to 16-year-old children with congenital hemiplegia: a systematic review. *Dev Med Child Neurol*. 2009.
5. Greaves S, Imms C, Dodd K, Krumlinde-Sundholm L. Assessing bimanual performance in young children with hemiplegic cerebral palsy: a systematic review. *Dev Med Child Neurol*. [Review]. 2010; **52**: 413-21.
6. Klingels K, Jaspers E, Van de Winckel A, De Cock P, Molenaers G, Feys H. A systematic review of arm activity measures for children with hemiplegic cerebral palsy. *Clin Rehabil*. 2010; **24**: 887-900.
7. Lemmens RJ, Timmermans AA, Janssen-Potten YJ, Smeets RJ, Seelen HA. Valid and reliable instruments for arm-hand assessment at ICF activity level in persons with hemiplegia: a systematic review. *BMC Neurol*. 2012; **12**: 21.
8. Krumlinde-Sundholm L, Ek L, Eliasson AC. What assessments evaluate use of hands in infants? A literature review. *Dev Med Child Neurol*. 2015; **57 Suppl 2**: 37-41.
9. Wallen M, Stewart K. Upper limb function in everyday life of children with cerebral palsy: description and review of parent report measures. *Disabil Rehabil*. 2014: 1-9.
10. Wagner LV, Davids JR. Assessment tools and classification systems used for the upper extremity in children with cerebral palsy. *Clin Orthop Relat Res*. 2012; **470**: 1257-71.
11. McConnell K, Johnston L, Kerr C. Upper limb function and deformity in cerebral palsy: a review of classification systems. *Dev Med Child Neurol*. 2011; **53**: 799-805.
12. Novak I, McIntyre S, Morgan C, et al. A systematic review of interventions for children with cerebral palsy: state of the evidence. *Dev Med Child Neurol*. 2013; **55**: 885-910.
13. Arner M, Eliasson AC, Nicklasson S, Sommerstein K, Hagglund G. Hand function in cerebral palsy. Report of 367 children in a population-based longitudinal health care program. *Journal of Hand Surgery (American Volume)*. 2008; **33**: 1337-47.
14. World Health Organization (WHO). International Classification of Functioning, Disability and Health. Geneva: WHO, 2001.
15. Krumlinde-Sundholm L, Eliasson AC. Development of the Assisting Hand Assessment: a Rasch-built measure intended for children with unilateral upper limb impairments. *Scandinavian Journal of Occupational Therapy*. 2003; **10**: 16-26.
16. World Health Organization. International classification of functioning, disability and health: ICF. Geneva: WHO2001.
17. Law M, McDermid J. Evidence-based rehabilitation : a guide to practice. 2nd ed. ed. Thorofare, N.J: Slack, 2008.
18. Cieza A, Geyh S, Chatterji S, Kostanjsek N, Ustun B, Stucki G. ICF linking rules: an update based on lessons learned. *Journal of Rehabilitation Medicine*. 2005; **37**: 212-8.
19. Portney LG, Watkins MP. Foundations of clinical research: applications to practice (3rd ed). New Jersey: Pearson Education, 2009.
20. Krumlinde-Sundholm L. Choosing and using assessments of hand function. In: Eliasson AC, Burtner, P.A., editor. Improving hand function in children with cerebral palsy: theory, evidence and intervention. London: Mac Keith Press, 2008. p. 176-97.
21. Terwee CB, Mokkink LB, Knol DL, Ostelo RW, Bouter LM, de Vet HC. Rating the methodological quality in systematic reviews of studies on measurement properties: a scoring system for the COSMIN checklist. *Qual Life Res*. 2012; **21**: 651-7.

22. Ammann-Reiffer C, Bastiaenen CH, de Bie RA, van Hedel HJ. Measurement properties of gait-related outcomes in youth with neuromuscular diagnoses: a systematic review. *Phys Ther*. 2014; **94**: 1067-82.
23. Gerber CN, Labruyere R, van Hedel HJ. Reliability and Responsiveness of Upper Limb Motor Assessments for Children With Central Neuromotor Disorders: A Systematic Review. *Neurorehabil Neural Repair*. 2015.
24. Terwee CB, Bot SD, de Boer MR, et al. Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol*. 2007; **60**: 34-42.
25. van Tulder M, Furlan A, Bombardier C, Bouter L. Updated method guidelines for systematic reviews in the cochrane collaboration back review group. *Spine (Phila Pa 1976)*. 2003; **28**: 1290-9.
26. Saether R, Helbostad JL, Riphagen, II, Vik T. Clinical tools to assess balance in children and adults with cerebral palsy: a systematic review. *Dev Med Child Neurol*. 2013; **55**: 988-99.
27. The ABILHAND-Kids questionnaire: a measure of manual ability in children with cerebral palsy. Bruxelles, Belgium: Université Catholique de Louvain. Available from: http://www.rehab-scales.org/abilhand-kids.html; [Last accessed 26 Nov 2015].
28. Erhardt RP. Developmental Hand Dysfunction: Theory-Assessment-Treatment: Laurel MD, Ramsco Publishing Company, 1982.
29. Folio MR, & Fewell, R.E. Peabody Developmental Motor Scales and Activity Cards. Chicago: Riverside, 1983.
30. DeMatteo C, Law M, Russell D, Pollock N, Rosenbaum P, Walter S. QUEST: Quality of upper extremity skills test manual. Hamilton, Ontario: Chedokee-McMaster Hospitals, 1992.
31. Randall M, Johnson LM, Reddihough DS. The Melbourne Assessment of Unilateral Upper Limb Function: test administration manual. Melbourne: Royal Children's Hospital, 1999.
32. Randall M, Johnson, L., Reddihough, D. The Melbourne Assessment 2: A test of Unilateral Upper Limb Function. Melbourne: The Royal Children's Hospital; 2012; Available from: http://www.rch.org.au/melbourneassessment/].
33. Folio MR, & Fewell, R.E. Peabody Developmental Motor Scales Second Edition. Austin, Texas: PRO-ED, 2000.
34. Mathiowetz V, Federman, S., Wiemer, D. Box and blocks test of manual dexterity: norms for 6-19 year olds. *Canadian Jounal of Occupational Therapy*. 1985; **52**: 241-5.
35. Jongbloed-Pereboom M, Nijhuis-van der Sanden MW, Steenbergen B. Norm scores of the box and block test for children ages 3-10 years. *Am J Occup Ther*. 2013; **67**: 312-8.
36. Taylor N, Sand PL, Jebsen RH. Evaluation of hand function in children. *Arch Phys Med Rehabil*. 1973; **54**: 129-35.
37. Bayley N. Bayley Scales of Infant and ToddlerDevelopmental Technical Manual. 3rd ed. San Antonio, Texas: Psychological Corporation, 2006.
38. Bayley N. Bayley Scales of Infant Development. 2nd ed. San Antonio, Texas: Psychological Coorporation, 1993.
39. Bruiniks RH. Bruiniks-Oseretsky Test of Motor Proficiency Examiner's Manual. Circle Pines, Minnesota: American Guidance Service, 1978.
40. Chien CW, Rodger S, Copley J. Development and Psychometric Evaluation of a New Measure for Children's Participation in Hand-Use Life Situations. *Arch Phys Med Rehabil*. 2015; **96**: 1045-55.
41. Erhardt RP. Developmental Hand Dysfunction: Theory, Assessment, and Treatment. 2nd ed. Austin, Texas: Pro-Ed, Inc., 1994.
42. Arnould C, Penta M, Renders A, Thonnard JL. ABILHAND-Kids: a measure of manual ability in children with cerebral palsy. *Neurology*. 2004; **63**: 1045-52.
43. Sorsdahl AB, Moe-Nilssen R, Strand LI. Observer reliability of the gross motor performance measure and the quality of upper extremity skills test, based on video recordings. *Dev Med Child Neurol*. 2008; **50**: 146-51.
44. Thorley M, Lannin N, Cusick A, Novak I, Boyd R. Construct validity of the Quality of Upper Extremity Skills Test for children with cerebral palsy. *Dev Med Child Neurol*. 2012; **54**: 1037-43.

45. Thorley M, Lannin N, Cusick A, Novak I, Boyd R. Reliability of the quality of upper extremity skills test for children with cerebral palsy aged 2 to 12 years. *Phys Occup Ther Pediatr*. 2012; **32**: 4-21.
46. Cusick A, Vasquez M, Knowles L, Wallen M. Effect of rater training on reliability of Melbourne Assessment of Unilateral Upper Limb Function scores. *Dev Med Child Neurol*. 2005; **47**: 39-45.
47. Bard R, Chaleat-Valayer E, Combey A, Bleu PE, Perretant I, Bernard JC. Upper limb assessment in children with cerebral palsy: Translation and reliability of the French version for the Melbourne unilateral upper limb assessment (test de Melbourne). *Annals of Physical and Rehabilitation Medicine*. 2009; **52**: 297-310.
48. Wang H, Liao H, Hsieh C. Reliability, sensitivity to change, and responsiveness of the Peabody Developmental Motor Scales-second edition for children with cerebral palsy. *Phys Ther*. 2006; **86**: 1351-9.
49. Randall M, Carlin JB, Chondros P, Reddihough D. Reliability of the Melbourne assessment of unilateral upper limb function. *Dev Med Child Neurol*. 2001; **43**: 761-7.
50. Randall M, Imms C, Carey L. Further evidence of validity of the Modified Melbourne Assessment for neurologically impaired children aged 2 to 4 years. *Dev Med Child Neurol*. 2012; **54**: 424-8.
51. Randall M, Imms C, Carey LM, Pallant JF. Rasch analysis of The Melbourne Assessment of Unilateral Upper Limb Function. *Dev Med Child Neurol*. 2014.
52. Johnson LM, Randall MJ, Reddihough DS, Oke LE, Byrt TA, Bach TM. Development of a clinical assessment of quality of movement for unilateral upper-limb functon. *Dev Med Child Neurol*. 1994; **36**: 965-73.
53. Bourke-Taylor H. Melbourne Assessment of Unilateral Upper Limb Function: construct validity and correlation with the Pediatric Evaluation of Disability Inventory. *Dev Med Child Neurol*. 2003; **45**: 92-6.
54. Erhardt RP, Beatty PA, Hertsgaard DM. A developmental prehension assessment for handicapped children. *Am J Occup Ther*. 1981; **35**: 237-42.
55. Russell DJ, Ward M, Law M. Test-retest reliability of the fine motor scale of the Peabody Developmental Motor Scales in children with cerebral palsy. *Occupational Therapy Journal of Research*. 1994; **14**: 178-82.
56. Marais I, Andrich D. Formalizing dimension and response violations of local independence in the unidimensional Rasch model. *Journal of Applied Measurement*. 2008; **9**: 200-15.
57. DeMatteo C, Law M, Russell D, Pollock N, Rosenbaum P, Walter S. The reliability and validity of the Quality of Upper Extremity Skills Test. *Physical and Occupational Therapy in Pediatrics*. 1993; **13**: 1-18.
58. Haga N, Hc, van Hoorn JF, Boonstra AM, Hadders-Algra M. Test-retest and inter- and intrareliability of the quality of the upper-extremity skills test in preschool-age children with cerebral palsy. *Arch Phys Med Rehabil*. 2007; **88**: 1686-9.
59. Mokkink LB, Terwee CB, Patrick DL, et al. The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. *J Clin Epidemiol*. 2010; **63**: 737-45.
60. Gao F, Mei X, Chen AC. Delayed finger tapping and cognitive responses in preterm-born male teenagers with mild spastic diplegia. *Pediatr Neurol*. 2015; **52**: 206-13.
61. Law M, Cadman D, Rosenbaum P, Walter S, Russell D, DeMatteo C. Neurodevelopmental therapy and upper-extremity inhibitive casting for children with cerebral palsy. *Dev Med Child Neurol*. 1991; **33**: 379-87.
62. Belvedere SL, de Morton NA. Application of Rasch analysis in health care is increasing and is applied for variable reasons in mobility instruments. *J Clin Epidemiol*. 2010; **63**: 1287-97.
63. Öhrvall AM, Krumlinde-Sundholm L, Eliasson AC. Exploration of the relationship between the Manual Ability Classification System and hand-function measures of capacity and performance. *Disabil Rehabil*. 2013; **35**: 913-8.
64. van Eck M, Dallmeijer AJ, van Lith IS, Voorman JM, Becher J. Manual ability and its relationship with daily activities in adolescents with cerebral palsy. *J Rehabil Med*. 2010; **42**: 493-8.

65. Holmefur M, Krumlinde-Sundholm L, Eliasson AC. Interrater and intrarater reliability of the Assisting Hand Assessment. *Am J Occup Ther*. 2007; **61**: 79-84.
66. Holmefur M, Krumlinde-Sundholm L. Psychometric properties of a revised version of the Assisting Hand Assessment (Kids-AHA 5.0). *Dev Med Child Neurol*. 2015.
67. Wallen M. Reflections on the contribution of the Assisting Hand Assessment. *Developmental Medicine & Childhood Neurology*. 2015.

Table I: Focus (according to ICF component and attribute measured), scale construction and standardization, as well as clinical utility of included hand function measures for children with bilateral cerebral palsy (CP)

Focus					Scale construction				Population			Clinical utility	
Outcome measure[a]	ICF component[b]	ICF qualifier	Primary purpose[c]	Attribute measured	Level[d]	Items[e]	Subscales[f]	Criterion or Norm	Age year	Target population	MACS GMFCS (I-V)[g]	Adm. time (min.)[h]	Examiner qualifications
EDPA[28]	A (59%)+BF	Capacity	E	Unimanual Reflexes Fine motor	O	n=17	3 subscales: • Reflexes • Voluntary movements • Pre-writing	Based on norms	0-6	CP or neurological condition		30-60	Recommended
PDMS[29] FMS	A (88%)+BF	Capacity	D,P,E	Fine motor	O	n=112	4 subscales: • Grasping • Hand use • Eye-hand coordination • Manual dexterity	Norm	0-7	Developmental delay		30	Required
PDMS-2[33] FMS	A (90%)+BF	Capacity	D,P,E	Fine motor	O	n=98	2 subscales: • Grasping • VMI	Norm	0-5	Developmental delay		30	Required
QUEST[30]	A (21%)+BF	Capacity	E	Unimanual Movement quality	N	n=33	4 subscales: • Dissociated movements • Grasp • Weight bearing • Protective extension	Criterion	2-12	CP	MACS I-V	30-45	Recommended

MA[31]	A (35%)+BF	Capacity	D,E	Unimanual Movement quality	O	n=16	One scale	Criterion	5-15	CP or neurological condition	GMFCS I-V	10-30	Required
MA2[32]	A (37%)+BF	Capacity	D,E	Movement quality Unimanual	O	n=14	4 subscales: • Range • Accuracy • Dexterity • Fluency	Criterion	2.5-15	CP or neurological condition	GMFCS I-V	10-30	Required
ABILHAND-Kids[27]	A (100%)	Performance	D	Perceived Manual ability	O	n=21	One scale Parent reported questionnaire		6-15	CP	GMFCS I-V	10-15	Not required

[a]EDPA= Erhardt Developmental Prehension Assessment; PDMS=Peabody Developmental Motor Scales; FMS=Fine Motor Scale; PDMS-2=Peabody Developmental Motor Scales Second Edition; QUEST=Quality of Upper Extremity Skills Test; MA= Melbourne Assessment of Unilateral Upper Limb Function; MA2=Melbourne Assessment 2, [b]A=activity (%=fine hand use); BF=body function, [c]D=discriminative; P=Predictive; E=Evaluative, [d]O=ordinal; N=nominal, [e]n=number, [f]VMI=Visual Motor Integration, [g]MACS=Manual Ability Classification System; GMFCS=Gross Motor Classification System, [h]Adm.=Administration; min.=minutes

Table II: The quality of studies assessing reliability of measures of hand function in children with bilateral cerebral palsy (CP) and the rating of the reported results of the studies

Outcome measure	Evaluated measurement property	N (Bilateral CP)	Age range; years	Results	Terwee Quality Rating	COSMIN Quality Rating	COSMIN "Worst score" item (excluding sample size and missing items)
EDPA	Inter-rater[54]	4 (4)	2.6-10.8	ICC=0.42-0.85	-/+	Fair	Minor methodological flaws
PDMS FMS	Test-retest[55]	18 (*)	1.8-7.9	$ICC_{1,1}$=0.99	+	Good	Assumable that measurements were independent
PDMS-2 FMS	Test-retest[48]	32 (27)	2.3-5.3	$ICC_{2,1}$=0.99	+	Good	Assumable that measurements were independent
	Measurement error[48]	32 (27)	2.3-5.3	SEM=1.3	+	Good	Assumable that measurements were independent
QUEST	Internal consistency[44]	94 (*) Grasp domain: 170 (74) assessments	2-16.6	*QUEST total score:* Not unidimensional *Unilateral grasp domain:* Unidimensional PSI=1.73-1.75; PSR=0.75	?	Poor	Important methodological flaws; repeated assessments; few items for grasp domain
	Internal consistency[45]	24 (11) 31 assessments	2.5-12.6	Cronbach alpha=0.97	+	Poor	Important methodological flaws; repeated assessments
	Intra-rater[58]	21 (15)	2-4.5	Spearman's rho (total score) =0.63-0.95 Spearman's rho (Grasp) =0.31-0.77	-/+	Fair	Statistical methods
	Intra-rater[43]	25 (21)	2-13	$ICC_{1,1}$ (total score)=0.69-0.89 $ICC_{1,1}$ (Grasp)=0.68-0.72	-/+	Excellent	
	Intra-rater[45]	24 (11) 31 assessments	2.5-12.6	ICC (total score)=0.96 ICC (Grasp)=0.90	+	Poor	Important methodological flaws; repeated assessments

	Inter-rater[57]	16 (*)	1.5-8	ICC (total score)=0.95	+	Fair	Minor methodological flaws; doubtful if raters assessed individually
	Inter-rater[57]	71 (43)	1.5-8	ICC (total score)=0.96 ICC (Grasp)=0.91	+	Fair	Minor methodological flaws; doubtful if raters assessed individually
	Inter-rater[57]	17 (*)	1.5-8	ICC (total score)=0.90 ICC (Grasp)=0.88	+	Fair	Minor methodological flaws; doubtful if raters assessed individually
	Inter-rater[58]	21 (15)	2-4.5	Spearman's rho(total score) =0.72-0.90 Spearman's rho (Grasp) =0.49-0.69	+	Fair	Statistical methods
	Inter-rater[43]	25 (21)	2-13	$ICC_{1,1}$ (total score)=0.91 $ICC_{1,1}$ (Grasp)=0.69	+	Excellent	
	Inter-rater[45]	24 (11) 31 assessments	2.5-12.6	ICC (total score):0.86 ICC (Grasp):0.67	+	Poor	Important methodological flaws; repeated assessments
	Test-retest[57]	17 (*)	1.5-8	ICC (total score)=0.95 ICC (Grasp)=0.93	+	Fair	Minor methodological flaws; doubtful if raters assessed individually
	Test-retest[58]	21 (15)	2-4.5	Spearman's rho(total score)=0.92 Spearman's rho (Grasp)=0.84	+	Fair	Statistical methods
MA	Internal consistency[49]	20 (20)	5-16	Cronbach alpha=0.96	+	Poor	No factor analysis
	Internal consistency[46]	9 (5)	5.5-12	Cronbach alpha=0.99	+	Poor	No factor analysis
	Intra-rater[52]	20 (20)	6-12	Kappa=0.79 and 0.82	+	Excellent	
	Intra-rater[49]	20 (20)	5-16	CCC=0.89-0.99 ICC=0.97	+	Excellent	
	Inter-rater[52]	20 (20)	6-12	Kappa=0.72	+	Excellent	
	Inter-rater[49]	20 (20)	5-16	ICC=0.95	+	Excellent	
	Inter-rater[46]	9 (5)	5.5-12	ICC=0.98- 0.99	+	Excellent	

	Test-retest[49]	20 (20)	5-16	CCC=0.97-0.98	+	Good	Assumable that patients were stable
	Measurement error[49]	20 (20)	5-16	Intra-rater: SDC: 14.3 Inter-rater: SDC: 17.2 Test-retest: Mean (LoA): 0.2 (-12.3, 12.6) and 2.4 (-10.2,15.1)	?	Excellent	
	Measurement error[46]	9 (5)	5.5-12	Inter-rater SEM: 2.56 (SD 26.7) and 3.37 (SD 25.1)	?	Excellent	
MA2	Internal consistency[51]	163 (86)	2-15	*Total score:* Not unidimensional *4 subscales:* Unidimensional PSI=0.81-0.92	 +	Excellent	
ABILHAND-Kids	Internal consistency[42]	113 (59)	6-15	Unidimensional scale PSR=0.94	+	Excellent	
	Test-retest[42]	113 (59)	6-15	Pearson's *r*=0.91	+	Good	Statistical methods

EDPA=Erhardt Developmental Prehension Assessment; PDMS=Peabody Developmental Motor Scales; FMS=Fine Motor Scales; PDMS-2= Peabody Developmental Motor Scales Second Edition; QUEST=Quality of Upper Extremity Skills Test; MA=Melbourne Assessment of unilateral upper limb function; MA2=Melbourne Assessment 2; N=Number; *=Children with bilateral CP included, but number unknown; SD=Standard deviation; PSR=Person separation reliability; ICC=Intra-class correlation; PSI=Person separation index; CCC=Lin concordance correlation coefficient; SDC=Smallest detectable difference; LoA=Limits of agreement; SEM=Standard error of measurement; +=positive; ?= indeterminate; -= negative

Table III: The quality of studies assessing validity and responsiveness of measures of hand function in children with bilateral cerebral palsy (CP) and the rating of the reported results of the studies

Outcome measure	Evaluated measurement property	N (Bilateral CP)	Age; Range years	Methods/Results	Terwee Quality Rating	COSMIN Quality Rating	COSMIN "Worst score" item (excluding sample size and missing items)
EDPA	Content validity[54]	4 (4)	2.6-10.8	Review of literature/Item judging by 16 OT's	+	Excellent	
QUEST	Content validity[57]	10 (*)	1.5-8	Review of literature and existing tools/Expert panel (n=4)	+	Excellent	
	Construct validity						
	Structural validity[44]	94 (*) Grasp domain: 170 (74) assessments	2-16.6	Rasch analysis Test for unidimensionality	?	Poor	Important methodological flaws; repeated assessments; few items for grasp domain
	Hypotheses testing[57]	71 (43)	1.5-8	Correlation between QUEST and PDMS FMS Pearson p=0.84	+	Fair	Some information on measurement properties of comparator instrument
MA	Content validity[52]	11 (11)	6-12	Review of literature and existing tools/Expert panel (n=4)	+	Excellent	
	Construct validity						
	Hypotheses testing[52]	11 (11)	6-12	Correlation between anticipated severity and MA: ICC=0.87	+	Poor	Measurement properties of comparator instrument
	Hypotheses testing[53]	18 (16)	5-14	Correlation between MA and PEDI Spearman's rho=0.72-0.94	+	Fair	Hypotheses vague
	Hypotheses testing[47]	46 (24)	5-16	Correlation between MA and Box and Blocks Pearson p=0.82	+	Poor	Unclear what was expected

MA2	Content validity[50]	30 (10)	2-4	Review of literature/Revision of items/Expert review	+	Excellent	
	Construct validity						
	Structural validity[51]	163 (86)	2-15	Rasch analysis Principal Component Analysis Test for unidimensionality	+	Excellent	
ABILHAND-Kids	Content validity[42]	113 (59)	6-15	Review of existing tools/Expert panel (n=27)	+	Excellent	
	Construct validity						
	Structural validity[42]	113 (59)	6-15	Rasch analysis	?	Excellent	
PDMS-2 FMS	Responsiveness[48]	32 (27)	2.3-5.3	GRI-R=1.7-2.3	+	Fair	Statistical methods applied not optimal

EDPA=Erhardt Developmental Prehension Assessment; PDMS=Peabody Developmental Motor Scales; FMS=Fine Motor Scales; PDMS-2= Peabody Developmental Motor Scales Second Edition; QUEST=Quality of Upper Extremity Skills Test; MA=Melbourne Assessment of unilateral upper limb function; MA2=Melbourne Assessment 2; N=Number; *=Children with bilateral CP included, but number unknown; SD=Standard deviation; ICC=Intra-class correlation; GRI-R=Guyatt responsive index for responsiveness; +=positive; ?= indeterminate; -= negative

Table IV: Reported measurement properties of the included hand function measures and their level of evidence (Table continues on next page)

	Outcome measures														
	EDPA			PDMS FMS			PDMS-2 FMS			QUEST			MA		
Measurement property	Ref	Level of evidence	TQR	Ref	Level of evidence	TQR	Ref	Level of evidence	TQR	Ref	Level of evidence	TQR	Ref	Level of evidence	TQR
Reliability															
Internal consistency															
Intra-rater										43,58	Limited	-/+-/+	49,52	Limited	++
Inter-rater	54	Unknown	-/+							43,57,58	Moderate	+++	49,52	Limited	+++
Test-retest				55	Unknown	+	48	Limited	+	57,58	Limited	++	49	Unknown	+
Measurement error							48	Limited	+				46,49	Unknown	??
Validity															
Content	54	Strong	+							57	Strong	+	52	Strong	+
Construct															
Structural validity															
Hypotheses testing										57	Limited	+			
Criterion															
Responsiveness							48	Limited	+						

EDPA=Erhardt Developmental Prehension Assessment; PDMS=Peabody Developmental Motor Scales; FMS=Fine Motor Scales; PDMS-2= Peabody Developmental Motor Scales Second Edition; QUEST=Quality of Upper Extremity Skills Test; MA=Melbourne Assessment of unilateral upper limb function; MA2=Melbourne Assessment 2; Ref=reference; TQR=Terwee Quality Rating; +=positive results; ?=indeterminate results; -/+=conflicting findings

Table IV continued

Measurement property	MA2 Ref	MA2 Level of evidence	MA2 TQR	ABILHAND-Kids Ref	ABILHAND-Kids Level of evidence	ABILHAND-Kids TQR
Reliability						
Internal consistency	51	Strong	+	42	Strong	+
Intra-rater						
Inter-rater						
Test-retest				42	Moderate	+
Measurement error						
Validity						
Content	50	Strong	+	42	Strong	+
Construct						
Structural validity	51	Strong	+	42	Strong	?
Hypotheses testing						
Criterion						
Responsiveness						

Figure legends

Figure 1: Processes performed to identify hand function measures and studies of measurement properties.

Acknowledgement

The study was funded by the Liaison Committee made up of representatives from the Central Norway Regional Health Authority (RHA) and the Norwegian University of Science and Technology (NTNU).

Supporting information

Additional supporting information may be found in the online version of this article:

Appendix S1: Database search.
Table S1: Quality criteria for measurement properties adapted from Terwee et al.[24]
Table S2: Levels of evidence for the overall quality of the measurement properties, based on the Cochrane Back Review Group.[25]
Table S3: Outcome measures used to measure hand function in children with cerebral palsy, but not included in the systematic review.

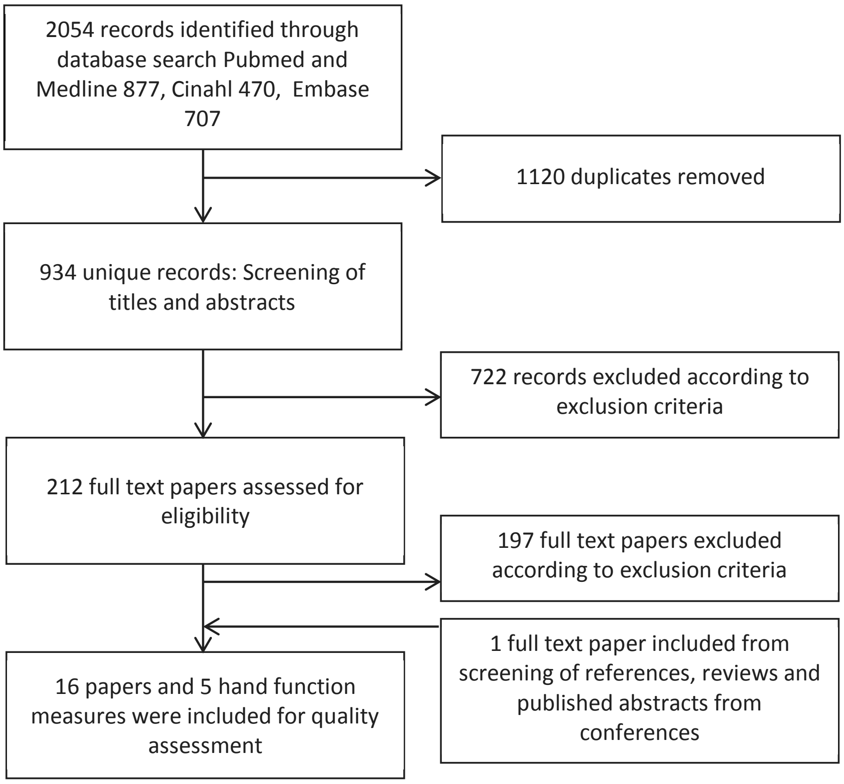


Figure 1: Processes performed to identify hand function measures and studies of measurement properties.

Appendix S1

Embase search through OvidSP (last update in edition 1980 to 2013 Week 22)

1a (*’cp’*)

cerebral palsy/

1b (*’diplegic cp’*)

Cerebral palsy/ and (diplegi* or bilateral OR quadriplegi* or dyskine* or ataxi*).ti,ab.

2a (*’hand function’ emtree*)

Hand function/ or Arm movement/ or ((arm/ or arm muscle/ or elbow/ or forearm/ or exp hand/ or wrist/) and (exp psychomotor activity/ or exp psychomotor performance/ or skill/ or motor activity/ or exp physical performance/[1] or musculoskeletal function/ or "movement (physiology)"/ or agility/ or exp limb movement/ or "range of motion"/ or voluntary movement/))

2b *(’hand function’ free text)*

((hand or hands or arm or arms or manual or bimanual or upper extremit* or upper limb*) adj3 (activit* or abilit* or agility or function* or perform* or skill*)).ti,ab.

3 (*’Instruments’*)

clinical assessment tool/ or scoring system/ or psychometry/ or measurement/ or rating scale/ or exp reliability/ or exp validity/ or "validation study"/

Final combination:

(1a and (2a or 2b) and 3) or (1b and (2a or 2b)) > 348

[1] Includes ’Motor performance’

Embase search 2 through OvidSP (last update in edition 1980 to 2014 Week 42)

1 (*'cp'*)

cerebral palsy/ or (cerebral pals* or little* disease or spastic diplegia* or spastic quadriplegia*).ti,ab.

2a *('names of tests and questionnaires' free text)*

(Melbourne Assessment of Unilateral Upper Limb Function).mp. or (Modified Melbourne Assessment or "Melbourne Assessment 2").mp. or (Quality of upper extremities skills test).mp. or (Jebsen-Taylor Test of hand function or Jebsen Test of Hand Function or Jebsen hand function test).mp. or (Peabody Developmental Motor Scale* adj3 (Second or "2") or prms2 or Peabody Developmental Fine Motor Scale*).mp. or (Bruininks-Oseretsky Test of Motor Proficiency).mp. or (Posture and Fine Motor Assessment of Infants and Fine motor part).mp. or (Bayley Scale* of Infant Development and Fine Motor Scale).mp. or (Griffiths Mental Developmental Scale* and Hand adj (Co-ordination or coordination)).mp. or (Denver Developmental Screening Tool* and fine motor part).mp. or (Erhardt Developmental Prehension Assessment).mp. or (ABILHAND-Kids) or (Assessment of children* hand skills).mp. or (Children* assessment of participation with hands).mp. or (CAP-Hand).mp. or (Children* hand-skills ability questionnaire).mp. or (In-Hand Manipulation Test or Exner In-Hand Manipulation Test or In-Hand Manipulation Skills or Test of In-Hand Manipulation).mp. or (Box and block* test).mp. or (Purdue Peg Board test or Purdue Pegboard Manual Dexterity Test).mp. or (Modified Sollermann* Grip Function Test or Sollerman hand function test).mp. or (Nine Hole Peg Test or "9-Hole Peg Test" or Peg Moving Task or Annett Peg Moving Task).mp. or (Computer Adaptive Test of Upper Extremity Function).mp. or (Upper Extremity Computer Adaptive Test) or (Developmental test of visual-motor integration).mp. or (Beery VMI).mp. or (Pediatric Outcomes Data Collection Instrument and Upper extremity).mp. or (Vineland Adaptive Behavior Scale* and fine motor part).mp. or (House functional classification or Consolidated House Classification).mp.

2b *('names of tests and questionnaires' in title)*

(Melbourne Assessment of Unilateral Upper Limb Function).ti. or (Modified Melbourne Assessment or "Melbourne Assessment 2").ti. or (Quality of upper extremities skills test).ti. or (Jebsen-Taylor Test of hand function or Jebsen Test of Hand Function or Jebsen hand function test).ti. or (Peabody Developmental Motor Scale* adj3 (Second or "2") or prms2 or Peabody Developmental Fine Motor Scale*).ti. or (Bruininks-Oseretsky Test of Motor Proficiency).ti. or (Posture and Fine Motor Assessment of Infants and Fine motor part).ti. or (Bayley Scale* of Infant Development and Fine Motor Scale).ti. or (Griffiths Mental Developmental Scale* and Hand adj (Co-ordination or coordination)).ti. or (Denver Developmental Screening Tool* and fine motor part).ti. or (Erhardt Developmental Prehension Assessment).ti. or (ABILHAND-Kids).ti. or (Assessment of children* hand skills).ti. or (Children* assessment of participation with hands).ti. or (CAP-Hand).ti. or (Children* hand-skills ability questionnaire).ti. or (In-Hand Manipulation Test or Exner In-Hand Manipulation Test or In-Hand Manipulation Skills or Test of In-Hand Manipulation).ti. or (Box and block* test).ti. or (Purdue Peg Board test or Purdue Pegboard Manual Dexterity Test).ti. or (Modified Sollermann* Grip Function Test or Sollerman hand function test).ti. or (Nine Hole Peg Test or "9-Hole Peg Test" or Peg

Moving Task or Annett Peg Moving Task).ti. or (Computer Adaptive Test of Upper Extremity Function).ti. or (Upper Extremity Computer Adaptive Test) or (Developmental test of visual-motor integration).ti. or (Beery VMI).ti. or (Pediatric Outcomes Data Collection Instrument and Upper extremity).ti. or (Vineland Adaptive Behavior Scale* and fine motor part).ti. or (House functional classification or Consolidated House Classification).ti.

3 (*’Instruments’*)

clinical assessment tool/ or scoring system/ or psychometry/ or measurement/ or rating scale/ or exp reliability/ or exp validity/ or "validation study"/

Final combination:

(1 and 2a and 3) or 2b > 190 hits

Embase update search through OvidSP (last update in edition 1980 to 2015 Week 24)

#	Searches	Results
1	cerebral palsy/	26538
2	Cerebral palsy/ and (diplegi* or bilateral or quadriplegi* or dyskine* or ataxi*).ti,ab.	3488
3	Hand function/ or Arm movement/ or ((arm/ or arm muscle/ or elbow/ or forearm/ or exp hand/ or wrist/) and (exp psychomotor activity/ or exp psychomotor performance/ or skill/ or motor activity/ or exp physical performance/ or musculoskeletal function/ or "movement (physiology)"/ or agility/ or exp limb movement/ or "range of motion"/ or voluntary movement/))	31948
4	((hand or hands or arm or arms or manual or bimanual or upper extremit* or upper limb*) adj3 (activit* or abilit* or agility or function* or perform* or skill*)).ti,ab.	27382
5	clinical assessment tool/ or scoring system/ or psychometry/ or measurement/ or rating scale/ or exp reliability/ or exp validity/ or "validation study"/	548979
6	(1 and (3 or 4) and 5) or (2 and (3 or 4))	487
7	6 and ("201323" or "201324" or "201325" or "201326" or "201327" or "201328" or "201329" or 20134* or 20135* or 2014* or 2015*).em.	142
8	cerebral palsy/ or (cerebral pals* or little* disease or spastic diplegia* or spastic quadriplegia*).ti,ab.	29278
9	(Melbourne Assessment of Unilateral Upper Limb Function or (Modified Melbourne Assessment or "Melbourne Assessment 2") or Quality of upper extremities skills test or (Jebsen-Taylor Test of hand function or Jebsen Test of Hand Function or Jebsen hand function test) or ((Peabody Developmental Motor Scale* adj3 (Second or "2")) or prms2 or Peabody Developmental Fine Motor Scale*) or Bruininks-Oseretsky Test of Motor Proficiency or (Posture and Fine Motor Assessment of Infants and Fine motor part) or (Bayley Scale* of Infant Development and Fine Motor Scale) or ((Griffiths Mental Developmental Scale* and Hand) adj (Co-ordination or coordination)) or (Denver Developmental Screening Tool* and fine motor part) or Erhardt Developmental Prehension Assessment or ABILHAND-Kids or Assessment of children* hand skills or Children* assessment of participation with hands or CAP-Hand or Children* hand-skills ability questionnaire or (In-Hand Manipulation Test or Exner In-Hand Manipulation Test or In-Hand Manipulation Skills or Test of In-Hand Manipulation) or (Box and block* test) or (Purdue Peg Board test or Purdue Pegboard Manual Dexterity Test) or (Modified Sollermann* Grip Function Test or Sollerman hand function test) or (Nine Hole Peg Test or "9-Hole Peg Test" or Peg Moving Task or Annett Peg Moving Task) or Computer Adaptive Test of Upper Extremity Function or Upper Extremity Computer Adaptive Test or Developmental test of visual-motor integration or Beery VMI or (Pediatric Outcomes Data Collection Instrument and Upper extremity) or (Vineland Adaptive Behavior Scale* and fine motor part) or (House functional classification or Consolidated House Classification)).mp. [mp=title,	1774

	abstract, subject headings, heading word, drug trade name, original title, device manufacturer, drug manufacturer, device trade name, keyword]	
10	(Melbourne Assessment of Unilateral Upper Limb Function or (Modified Melbourne Assessment or "Melbourne Assessment 2") or Quality of upper extremities skills test or (Jebsen-Taylor Test of hand function or Jebsen Test of Hand Function or Jebsen hand function test) or ((Peabody Developmental Motor Scale* adj3 (Second or "2")) or prms2 or Peabody Developmental Fine Motor Scale*) or Bruininks-Oseretsky Test of Motor Proficiency or (Posture and Fine Motor Assessment of Infants and Fine motor part) or (Bayley Scale* of Infant Development and Fine Motor Scale) or ((Griffiths Mental Developmental Scale* and Hand) adj (Co-ordination or coordination)) or (Denver Developmental Screening Tool* and fine motor part) or Erhardt Developmental Prehension Assessment or ABILHAND-Kids or Assessment of children* hand skills or Children* assessment of participation with hands or CAP-Hand or Children* hand-skills ability questionnaire or (In-Hand Manipulation Test or Exner In-Hand Manipulation Test or In-Hand Manipulation Skills or Test of In-Hand Manipulation) or (Box and block* test) or (Purdue Peg Board test or Purdue Pegboard Manual Dexterity Test) or (Modified Sollermann* Grip Function Test or Sollerman hand function test) or (Nine Hole Peg Test or "9-Hole Peg Test" or Peg Moving Task or Annett Peg Moving Task) or Computer Adaptive Test of Upper Extremity Function or Upper Extremity Computer Adaptive Test or Developmental test of visual-motor integration or Beery VMI or (Pediatric Outcomes Data Collection Instrument and Upper extremity) or (Vineland Adaptive Behavior Scale* and fine motor part) or (House functional classification or Consolidated House Classification)).ti.	132
11	clinical assessment tool/ or scoring system/ or psychometry/ or measurement/ or rating scale/ or exp reliability/ or exp validity/ or "validation study"/	548979
12	(8 and 9 and 11) or 10	207
13	12 and ("201443" or "201444" or "201445" or "201446" or "201447" or "201448" or "201449" or 20145* or 2015*).em.	19
14	7 or 13	150

6 or 12 (total number of hits): 625

MEDLINE search through OvidSP (last update in edition 1946 to 2013 May Week 4)

1a (*'cp'*)

Cerebral Palsy/ OR (cerebral pals* OR little* disease OR spastic diplegia* OR spastic quadriplegia*).ti,ab.

1b (*'diplegic cp'*)

(Cerebral Palsy/ OR (cerebral pals* OR little* disease).ti,ab.) AND (diplegi* OR bilateral OR quadriplegi* OR dyskine* OR ataxi*).ti,ab.

2a *('hand function' MeSH)*

Exp Upper Extremity/ph, pp OR exp Hand Strength/ OR (exp Upper Extremity/ AND (exp "Musculoskeletal Physiological Phenomena"/ OR exp "Motor Skills"/ OR exp "Task Performance and Analysis"/))

2b *('hand function' free text)*

((hand or hands or arm or arms or manual or bimanual or upper extremit* or upper limb*) adj3 (activit* or abilit* or agility or function* or perform* or skill*)).ti,ab.

3a (*'Instruments'*)

(Analy* OR evaluat*).ti. OR (assessment* OR assessing OR instrument OR instruments OR measure OR measurement* OR measures OR quantifying OR quantification OR questionnaire* OR scale OR scales OR score OR scores OR screening OR subtest* OR test OR tests OR testing OR tool OR tools).ti,ab.

3b (*'measurement properties'*)

Disability evaluation/ OR Evaluation Studies.pt. OR Observer variation/ OR Psychometrics/ OR Reference Values/ OR exp "Reproducibility of Results"/ OR exp "Sensitivity and Specificity"/ OR Validation studies.pt. OR (accura* OR clinimetr* OR coefficient* OR consisten* OR correlate* OR cronbach OR discrimina* OR feasib* OR interrater OR inter-rater OR intersession OR inter-session OR intertester OR inter-tester OR intrarater OR intra-rater OR intratester OR intra-tester OR kappa OR predictiv* OR propert* OR psychometr* OR reliab* OR repeatab* OR reproducib* OR responsive* OR sensitiv* OR spearman* OR specific* OR spearman OR subscale* OR suitab* OR test development OR test-retest OR useful* OR utility OR valid* OR variance).ti,ab.

Final combination:

(1a AND (2a OR 2b) AND (3a OR 3b)) OR (1b AND (2a OR 2b)) > 565 hits

MEDLINE search 2 through OvidSP (last update in edition 1946 to 2014 October week 3)

1 (*'cp'*)

Cerebral Palsy/ OR (cerebral pals* OR little* disease OR spastic diplegia* OR spastic quadriplegia*).ti,ab.

2a *('names of tests and questionnaires' free text)*

(Melbourne Assessment of Unilateral Upper Limb Function).mp. or (Modified Melbourne Assessment or "Melbourne Assessment 2").mp. or (Quality of upper extremities skills test).mp. or (Jebsen-Taylor Test of hand function or Jebsen Test of Hand Function or Jebsen hand function test).mp. or (Peabody Developmental Motor Scale* adj3 (Second or "2") or prms2 or Peabody Developmental Fine Motor Scale*).mp. or (Bruininks-Oseretsky Test of Motor Proficiency).mp. or (Posture and Fine Motor Assessment of Infants and Fine motor part).mp. or (Bayley Scale* of Infant Development and Fine Motor Scale).mp. or (Griffiths Mental Developmental Scale* and Hand adj (Co-ordination or coordination)).mp. or (Denver Developmental Screening Tool* and fine motor part).mp. or (Erhardt Developmental Prehension Assessment).mp. or (ABILHAND-Kids).mp. or (Assessment of children* hand skills).mp. or (Children* assessment of participation with hands).mp. or (CAP-Hand).mp. or (Children* hand-skills ability questionnaire).mp. or (In-Hand Manipulation Test or Exner In-Hand Manipulation Test or In-Hand Manipulation Skills or Test of In-Hand Manipulation).mp. or (Box and block* test).mp. or (Purdue Peg Board test or Purdue Pegboard Manual Dexterity Test).mp. or (Modified Sollermann* Grip Function Test or Sollerman hand function test).mp. or (Nine Hole Peg Test or "9-Hole Peg Test" or Peg Moving Task or Annett Peg Moving Task).mp. or (Computer Adaptive Test of Upper Extremity Function).mp. or (Upper Extremity Computer Adaptive Test) or (Developmental test of visual-motor integration).mp. or (Beery VMI).mp. or (Pediatric Outcomes Data Collection Instrument and Upper extremity).mp. or (Vineland Adaptive Behavior Scale* and fine motor part).mp. or (House functional classification or Consolidated House Classification).mp.

2b *('names of tests and questionnaires' in title)*

(Melbourne Assessment of Unilateral Upper Limb Function).ti. or (Modified Melbourne Assessment or "Melbourne Assessment 2").ti. or (Quality of upper extremities skills test).ti. or (Jebsen-Taylor Test of hand function or Jebsen Test of Hand Function or Jebsen hand function test).ti. or (Peabody Developmental Motor Scale* adj3 (Second or "2") or prms2 or Peabody Developmental Fine Motor Scale*).ti. or (Bruininks-Oseretsky Test of Motor Proficiency).ti. or (Posture and Fine Motor Assessment of Infants and Fine motor part).ti. or (Bayley Scale* of Infant Development and Fine Motor Scale).ti. or (Griffiths Mental Developmental Scale* and Hand adj (Co-ordination or coordination)).ti. or (Denver Developmental Screening Tool* and fine motor part).ti. or (Erhardt Developmental Prehension Assessment).ti. or (ABILHAND-Kids).ti. or (Assessment of children* hand skills).ti. or (Children* assessment of participation with hands).ti. or (CAP-Hand).ti. or (Children* hand-skills ability questionnaire).ti. or (In-Hand Manipulation Test or Exner In-Hand Manipulation Test or In-Hand Manipulation Skills or Test of In-Hand Manipulation).ti. or (Box and block* test).ti. or (Purdue Peg Board test or Purdue Pegboard Manual Dexterity Test).ti. or (Modified Sollermann* Grip

Function Test or Sollerman hand function test).ti. or (Nine Hole Peg Test or "9-Hole Peg Test" or Peg Moving Task or Annett Peg Moving Task).ti. or (Computer Adaptive Test of Upper Extremity Function).ti. or (Upper Extremity Computer Adaptive Test) or (Developmental test of visual-motor integration).ti. or (Beery VMI).ti. or (Pediatric Outcomes Data Collection Instrument and Upper extremity).ti. or (Vineland Adaptive Behavior Scale* and fine motor part).ti. or (House functional classification or Consolidated House Classification).ti.

3 (*'measurement properties'*)

Disability evaluation/ OR Evaluation Studies.pt. OR Observer variation/ OR Psychometrics/ OR Reference Values/ OR exp "Reproducibility of Results"/ OR exp "Sensitivity and Specificity"/ OR Validation studies.pt. OR (accura* OR clinimetr* OR coefficient* OR consisten* OR correlate* OR cronbach OR discrimina* OR feasib* OR interrater OR inter-rater OR intersession OR inter-session OR intertester OR inter-tester OR intrarater OR intra-rater OR intratester OR intra-tester OR kappa OR predictiv* OR propert* OR psychometr* OR reliab* OR repeatab* OR reproducib* OR responsive* OR sensitiv* OR spearman* OR specific* OR spearman OR subscale* OR suitab* OR test development OR test-retest OR useful* OR utility OR valid* OR variance).ti,ab.

<u>Final combination:</u>

(1 and 2a and 3) or 2b > 178 hits

MEDLINE update search through OvidSP (last update in edition 1946 to 2015 June week 1)

#	Searches	Results
1	Cerebral Palsy/ or (cerebral pals* or little* disease or spastic diplegia* or spastic quadriplegia*).ti,ab.	19833
2	(Cerebral Palsy/ or (cerebral pals* or little* disease).ti,ab.) and (diplegi* or bilateral or quadriplegi* or dyskine* or ataxi*).ti,ab.	2396
3	exp Upper Extremity/ph, pp or exp Hand Strength/ or (exp Upper Extremity/ and (exp "Musculoskeletal Physiological Phenomena"/ or exp "Motor Skills"/ or exp "Task Performance and Analysis"/))	46833
4	((hand or hands or arm or arms or manual or bimanual or upper extremit* or upper limb*) adj3 (activit* or abilit* or agility or function* or perform* or skill*)).ti,ab.	19184
5	(Analy* or evaluat*).ti. or (assessment* or assessing or instrument or instruments or measure or measurement* or measures or quantifying or quantification or questionnaire* or scale or scales or score or scores or screening or subtest* or test or tests or testing or tool or tools).ti,ab.	4731086
6	Disability evaluation/ or Evaluation Studies.pt. or Observer variation/ or Psychometrics/ or Reference Values/ or exp "Reproducibility of Results"/ or exp "Sensitivity and Specificity"/ or Validation studies.pt. or (accura* or clinimetr* or coefficient* or consisten* or correlate* or cronbach or discrimina* or feasib* or interrater or inter-rater or intersession or inter-session or intertester or inter-tester or intrarater or intra-rater or intratester or intra-tester or kappa or predictiv* or propert* or psychometr* or reliab* or repeatab* or reproducib* or responsive* or sensitiv* or spearman* or specific* or spearman or subscale* or suitab* or test development or test-retest or useful* or utility or valid* or variance).ti,ab.	6127629
7	(1 and (3 or 4) and (5 or 6)) or (2 and (3 or 4))	700
8	7 and (2013052* or 201306* or 201307* or 201308* or 201309* or 20131* or 2014* or 2015*).ed.	143
9	(Melbourne Assessment of Unilateral Upper Limb Function or (Modified Melbourne Assessment or "Melbourne Assessment 2") or Quality of upper extremities skills test or (Jebsen-Taylor Test of hand function or Jebsen Test of Hand Function or Jebsen hand function test) or ((Peabody Developmental Motor Scale* adj3 (Second or "2")) or prms2 or Peabody Developmental Fine Motor Scale*) or Bruininks-Oseretsky Test of Motor Proficiency or (Posture and Fine Motor Assessment of Infants and Fine motor part) or (Bayley Scale* of Infant Development and Fine Motor Scale) or ((Griffiths Mental Developmental Scale* and Hand) adj (Co-ordination or coordination)) or (Denver Developmental Screening Tool* and fine motor part) or Erhardt Developmental Prehension Assessment or ABILHAND-Kids or Assessment of children* hand skills or Children* assessment of participation with hands or CAP-Hand or Children* hand-skills ability questionnaire or (In-Hand	972

	Manipulation Test or Exner In-Hand Manipulation Test or In-Hand Manipulation Skills or Test of In-Hand Manipulation) or (Box and block* test) or (Purdue Peg Board test or Purdue Pegboard Manual Dexterity Test) or (Modified Sollermann* Grip Function Test or Sollerman hand function test) or (Nine Hole Peg Test or "9-Hole Peg Test" or Peg Moving Task or Annett Peg Moving Task) or Computer Adaptive Test of Upper Extremity Function or Upper Extremity Computer Adaptive Test or Developmental test of visual-motor integration or Beery VMI or (Pediatric Outcomes Data Collection Instrument and Upper extremity) or (Vineland Adaptive Behavior Scale* and fine motor part) or (House functional classification or Consolidated House Classification)).mp. [mp=title, abstract, original title, name of substance word, subject heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier]	
10	(Melbourne Assessment of Unilateral Upper Limb Function or (Modified Melbourne Assessment or "Melbourne Assessment 2") or Quality of upper extremities skills test or (Jebsen-Taylor Test of hand function or Jebsen Test of Hand Function or Jebsen hand function test) or ((Peabody Developmental Motor Scale* adj3 (Second or "2")) or prms2 or Peabody Developmental Fine Motor Scale*) or Bruininks-Oseretsky Test of Motor Proficiency or (Posture and Fine Motor Assessment of Infants and Fine motor part) or (Bayley Scale* of Infant Development and Fine Motor Scale) or ((Griffiths Mental Developmental Scale* and Hand) adj (Co-ordination or coordination)) or (Denver Developmental Screening Tool* and fine motor part) or Erhardt Developmental Prehension Assessment or ABILHAND-Kids or Assessment of children* hand skills or Children* assessment of participation with hands or CAP-Hand or Children* hand-skills ability questionnaire or (In-Hand Manipulation Test or Exner In-Hand Manipulation Test or In-Hand Manipulation Skills or Test of In-Hand Manipulation) or (Box and block* test) or (Purdue Peg Board test or Purdue Pegboard Manual Dexterity Test) or (Modified Sollermann* Grip Function Test or Sollerman hand function test) or (Nine Hole Peg Test or "9-Hole Peg Test" or Peg Moving Task or Annett Peg Moving Task) or Computer Adaptive Test of Upper Extremity Function or Upper Extremity Computer Adaptive Test or Developmental test of visual-motor integration or Beery VMI or (Pediatric Outcomes Data Collection Instrument and Upper extremity) or (Vineland Adaptive Behavior Scale* and fine motor part) or (House functional classification or Consolidated House Classification)).ti.	104
11	Disability evaluation/ or Evaluation Studies.pt. or Observer variation/ or Psychometrics/ or Reference Values/ or exp "Reproducibility of Results"/ or exp "Sensitivity and Specificity"/ or Validation studies.pt. or (accura* or clinimetr* or coefficient* or consisten* or correlate* or cronbach or discrimina* or feasib* or interrater or inter-rater or intersession or inter-session or intertester or inter-tester or intrarater or intra-rater or intratester or intra-tester or kappa or predictiv* or propert* or psychometr* or reliab* or repeatab* or reproducib* or responsive* or sensitiv* or spearman* or specific* or spearman or subscale* or suitab* or test development or test-retest or useful* or utility or valid* or variance).ti,ab.	6127629

12	(1 and 9 and 11) or 10	174 *
13	12 and (2014102* or 201411* or 201412* or 2015*).ed.	4
14	8 or 13	144

7 or 12 > 803 hits

CINAHL search through EbscoHost (last update in edition June 2013)

1a (*’cp broad’*)

MH "Cerebral Palsy" OR TI "cerebral pals*" OR AB "cerebral pals*" OR TI "little* disease" OR AB "little* disease" OR TI "spastic diplegia*" OR AB "spastic diplegia*" OR TI "spastic quadriplegia*" OR AB "spastic quadriplegia*"

1b (*’cp diplegic’*)

((MH "Cerebral Palsy" OR TI "cerebral pals*" OR AB "cerebral pals*" OR TI "little* disease" OR AB "little* disease") AND (TI diplegi* OR TI bilateral OR TI quadriplegi* OR TI dyskine* OR TI ataxi* OR AB diplegi* OR AB bilateral OR AB quadriplegi* OR AB dyskine* OR AB ataxi*)) OR TI "spastic diplegia*" OR AB "spastic diplegia*" OR TI "spastic quadriplegia*" OR AB "spastic quadriplegia*"

2a (*’hand function’ subject headings*)

((MH "Upper Extremity+") AND (MH "Musculoskeletal System Physiology+" OR MH "Motor Skills+")) OR IN manual OR IN bimanual OR IN hand OR IN hands OR IN arm OR IN arms OR IN upper limb* OR IN upper extremit*

2b (*’hand function’ free text*)

((TI hand or TI hands or TI arm or TI arms or TI manual or TI bimanual or TI upper extremit* or TI upper limb*) N3 (TI activit* or TI abilit* or TI agility or TI function* or TI perform* or TI skill*)) OR ((AB hand or AB hands or AB arm or AB arms or AB manual or AB bimanual or AB upper extremit* or AB upper limb*) N3 (AB activit* or AB abilit* or AB agility or AB function* or AB perform* or AB skill*))

3 (*’instruments/properties’*)

MH "Instrument Validation" OR MH "Clinical Assessment Tools" OR MH "Questionnaires" OR MH "Research Instruments" OR MH "Reproducibility of Results" OR MH "Validation Studies" OR MH "Test-Retest Reliability" OR MH "Pearson's Correlation Coefficient" OR MH "Spearman's Rank Correlation Coefficient" OR MH "Concurrent Validity" OR MH "Construct Validity" OR MH "Intraclass Correlation Coefficient" OR MH "Interrater Reliability" OR MH "Intrarater Reliability" OR MH "Kappa Statistic" OR MH "Instrument Construction" OR IN manual OR IN bimanual OR IN hand OR IN hands OR IN arm OR IN arms OR IN upper limb* OR IN upper extremit*

Final combination:

(1a AND (2a OR 2b) AND 3) OR (1b AND (2a OR 2b)) >251 hits

CINAHL search 2 through EbscoHost (last update in edition ---- 2014)

1 (*’cp’*)

MH "Cerebral Palsy" OR TI "cerebral pals*" OR AB "cerebral pals*" OR TI "little* disease" OR AB "little* disease" OR TI "spastic diplegia*" OR AB "spastic diplegia*" OR TI "spastic quadriplegia*" OR AB "spastic quadriplegia*"

2a *(‘names of tests and questionnaires’ free text)*

TX ("Melbourne Assessment of Unilateral Upper Limb Function") or TX ("Modified Melbourne Assessment" or "Melbourne Assessment 2") or TX ("Quality of upper extremities skills test") or TX ("Jebsen-Taylor Test of hand function" or "Jebsen Test of Hand Function" or "Jebsen hand function test") or TX ("Peabody Developmental Motor Scale*" W3 (Second or "2") or prms2 or "Peabody Developmental Fine Motor Scale*") or TX ("Bruininks-Oseretsky Test of Motor Proficiency") or TX ("Posture and Fine Motor Assessment of Infants" and "Fine motor part") or TX ("Bayley Scale* of Infant Development" and "Fine Motor Scale") or TX ("Griffiths Mental Developmental Scale*" and Hand W2 (Co-ordination or coordination)) or TX ("Denver Developmental Screening Tool*" and "fine motor part") or TX ("Erhardt Developmental Prehension Assessment") or TX (ABILHAND-Kids) or TX ("Assessment of children* hand skills") or TX ("Children* assessment of participation with hands") or TX (CAP-Hand) or TX ("Children* hand-skills ability questionnaire") or TX ("In-Hand Manipulation Test" or "Exner In-Hand Manipulation Test" or "In-Hand Manipulation Skills" or "Test of In-Hand Manipulation") or TX ("Box and block* test") or TX ("Purdue Peg Board test" or "Purdue Pegboard Manual Dexterity Test") or TX ("Modified Sollermann* Grip Function Test" or "Sollerman hand function test") or TX ("Nine Hole Peg Test" or "9-Hole Peg Test" or "Peg Moving Task" or "Annett Peg Moving Task") or TX ("Computer Adaptive Test of Upper Extremity Function") or TX ("Upper Extremity Computer Adaptive Test") or TX ("Developmental test of visual-motor integration") or TX ("Beery VMI") or TX ("Pediatric Outcomes Data Collection Instrument" and "Upper extremity") or TX ("Vineland Adaptive Behavior Scale*" and "fine motor part") or TX ("House functional classification" or "Consolidated House Classification")

2b *(‘names of tests and questionnaires’ in title)*

TI (Melbourne Assessment of Unilateral Upper Limb Function) or TI (Modified Melbourne Assessment or "Melbourne Assessment 2") or TI (Quality of upper extremities skills test) or TI (Jebsen-Taylor Test of hand function or Jebsen Test of Hand Function or Jebsen hand function test) or TI (Peabody Developmental Motor Scale* W3 (Second or "2") or prms2 or Peabody Developmental Fine Motor Scale*) or TI (Bruininks-Oseretsky Test of Motor Proficiency) or TI (Posture and Fine Motor Assessment of Infants and Fine motor part) or TI (Bayley Scale* of Infant Development and Fine Motor Scale) or TI (Griffiths Mental Developmental Scale* and Hand W2 (Co-ordination or coordination)) or TI (Denver Developmental Screening Tool* and fine motor part) or TI (Erhardt Developmental Prehension Assessment) or TI (ABILHAND-Kids) or TI (Assessment of children* hand skills) or TI (Children* assessment of participation with hands) or TI (CAP-Hand) or TI (Children* hand-skills ability questionnaire) or TI (In-Hand Manipulation Test or Exner In-Hand Manipulation Test or

In-Hand Manipulation Skills or Test of In-Hand Manipulation) or TI (Box and block* test) or TI (Purdue Peg Board test or Purdue Pegboard Manual Dexterity Test) or TI (Modified Sollermann* Grip Function Test or Sollerman hand function test) or TI (Nine Hole Peg Test or "9-Hole Peg Test" or Peg Moving Task or Annett Peg Moving Task) or TI (Computer Adaptive Test of Upper Extremity Function) or TI (Upper Extremity Computer Adaptive Test) or TI (Developmental test of visual-motor integration) or TI (Beery VMI) or TI (Pediatric Outcomes Data Collection Instrument and Upper extremity) or TI (Vineland Adaptive Behavior Scale* and fine motor part) or TI (House functional classification or Consolidated House Classification)

3 (*'instruments/properties'*)

MH "Instrument Validation" OR MH "Clinical Assessment Tools" OR MH "Questionnaires" OR MH "Research Instruments" OR MH "Reproducibility of Results" OR MH "Validation Studies" OR MH "Test-Retest Reliability" OR MH "Pearson's Correlation Coefficient" OR MH "Spearman's Rank Correlation Coefficient" OR MH "Concurrent Validity" OR MH "Construct Validity" OR MH "Intraclass Correlation Coefficient" OR MH "Interrater Reliability" OR MH "Intrarater Reliability" OR MH "Kappa Statistic" OR MH "Instrument Construction" OR IN manual OR IN bimanual OR IN hand OR IN hands OR IN arm OR IN arms OR IN upper limb* OR IN upper extremit*

<u>Final combination:</u>

(1 AND 2a AND 3) OR 2b > 175 hits

CINAHL search through EbscoHost (last update in edition (10) June 2015)

Search ID#	Search Terms	Actions
S12	S6 OR S10 Published Date: 20130101-20151231	(73)
S11	S6 OR S10	(410)
S10	(s1 AND s7 AND s9) OR s8	(185)
S9	MH "Instrument Validation" OR MH "Clinical Assessment Tools" OR MH "Questionnaires" OR MH "Research Instruments" OR MH "Reproducibility of Results" OR MH "Validation Studies" OR MH "Test-Retest Reliability" OR MH "Pearson's Correlation Coefficient" OR MH "Spearman's Rank Correlation Coefficient" OR MH "Concurrent Validity" OR MH "Construct Validity" OR MH "Intraclass Correlation Coefficient" OR MH "Interrater Reliability" OR MH "Intrarater Reliability" OR MH "Kappa Statistic" OR MH "Instrument C ...	(327,707)
S8	TI (Melbourne Assessment of Unilateral Upper Limb Function) or TI (Modified Melbourne Assessment or "Melbourne Assessment 2") or TI (Quality of upper extremities skills test) or TI (Jebsen-Taylor Test of hand function or Jebsen Test of Hand Function or Jebsen hand function test) or TI (Peabody Developmental Motor Scale* W3 (Second or "2") or prms2 or Peabody Developmental Fine Motor Scale*) or TI (Bruininks-Oseretsky Test of Motor Proficiency) or TI (Posture and Fine Motor Assessment of Infants ...	(94)
S7	TX ("Melbourne Assessment of Unilateral Upper Limb Function") or TX ("Modified Melbourne Assessment" or "Melbourne Assessment 2") or TX ("Quality of upper extremities skills test") or TX ("Jebsen-Taylor Test of hand function" or "Jebsen Test of Hand Function" or "Jebsen hand function test") or TX ("Peabody Developmental Motor Scale*" W3 (Second or "2") or prms2 or "Peabody Developmental Fine Motor Scale*") or TX ("Bruininks-Oseretsky Test of Motor Proficiency") or TX ("Posture and Fine Motor Ass ...	(1,034)
S6	(s1 AND (s3 OR s4) AND s5) OR (s2 AND (s3 OR s4))	(298)
S5	MH "Instrument Validation" OR MH "Clinical Assessment Tools" OR MH "Questionnaires" OR MH "Research Instruments" OR MH "Reproducibility of Results" OR MH "Validation Studies" OR MH "Test-Retest Reliability" OR MH "Pearson's Correlation Coefficient" OR MH "Spearman's Rank Correlation Coefficient" OR MH "Concurrent Validity" OR MH "Construct Validity" OR MH "Intraclass Correlation	(327,707)

	Coefficient" OR MH "Interrater Reliability" OR MH "Intrarater Reliability" OR MH "Kappa Statistic" OR MH "Instrument C ...	
S4	((TI hand or TI hands or TI arm or TI arms or TI manual or TI bimanual or TI upper extremit* or TI upper limb*) N3 (TI activit* or TI abilit* or TI agility or TI function* or TI perform* or TI skill*)) OR ((AB hand or AB hands or AB arm or AB arms or AB manual or AB bimanual or AB upper extremit* or AB upper limb*) N3 (AB activit* or AB abilit* or AB agility or AB function* or AB perform* or AB skill*))	(5,430)
S3	((MH "Upper Extremity+") AND (MH "Musculoskeletal System Physiology+" OR MH "Motor Skills+")) OR IN manual OR IN bimanual OR IN hand OR IN hands OR IN arm OR IN arms OR IN upper limb* OR IN upper extremit*	(8,972)
S2	((MH "Cerebral Palsy" OR TI "cerebral pals*" OR AB "cerebral pals*" OR TI "little* disease" OR AB "little* disease") AND (TI diplegi* OR TI bilateral OR TI quadriplegi* OR TI dyskine* OR TI ataxi* OR AB diplegi* OR AB bilateral OR AB quadriplegi* OR AB dyskine* OR AB ataxi*)) OR TI "spastic diplegia*" OR AB "spastic diplegia*" OR TI "spastic quadriplegia*" OR AB "spastic quadriplegia*"	(830)
S1	MH "Cerebral Palsy" OR TI "cerebral pals*" OR AB "cerebral pals*" OR TI "little* disease" OR AB "little* disease" OR TI "spastic diplegia*" OR AB "spastic diplegia*" OR TI "spastic quadriplegia*" OR AB "spastic quadriplegia*"	(6,999)

Table S1: Quality criteria for measurement properties adapted from Terwee et al (2007)

Property	Rating	Quality Criteria
Reliability		
Internal consistency	+	(Sub)scale unidimensional AND Cronbach's alpha(s) ≥0.70
	?	Dimensionality not known OR Cronbach's alpha not determined
	-	(Sub)scale not unidimensional OR Cronbach's alpha(s) <0.70
Reliability	+	ICC/weighted Kappa ≥0.70 OR Pearson's *r* ≥0.80
	?	Neither ICC/weighted Kappa, nor Pearson's *r* determined
	-	ICC/weighted Kappa <0.70 OR Pearson's *r* <0.80
Measurement error	+	MIC > SDC OR MIC outside the LOA
	?	MIC not defined
	-	MIC ≤ SDC OR MIC equals or inside the LOA
Validity		
Content validity	+	A clear description is provided of the measurement aim, the target population, the concept(s) being measured, and the item selection. Target population and experts in the field were involved in the development process
	?	Not enough information available
	-	No clear description is provided of the measurement aim, the target population, the concept(s) being measured, and the item selection, OR target population and experts in the field were not involved in the developmental process
Construct validity		
Structural validity	+	Factors should explain ≥50% of the variance
	?	Explained variance not mentioned
	-	Factors explain <50% of the variance
Hypothesis testing	+	Correlations with instruments measuring the same construct ≥0.50 OR ≥75% of the results are in accordance with the hypotheses AND correlations with related constructs are higher than with unrelated constructs
	?	Solely correlations determined with unrelated constructs
	-	Correlations with instruments measuring the same construct <0.50 OR <75% of the results are in accordance with the hypotheses AND correlations with related constructs are lower than with unrelated constructs
Cross-cultural validity	+	No differences in factor structure OR no important DIF between language versions
	?	Multiple group factor analysis not applied AND DIF not assessed
	-	Differences in factor structure OR important DIF between language versions
Criterion validity	+	Convincing arguments that gold standard is "gold" AND correlation with gold standard ≥0.70
	?	No convincing arguments that gold standard is "gold" OR doubtful design or methods
	-	Correlation with gold standard <0.70 despite adequate design and method
Responsiveness		
Responsiveness	+	Correlation with changes on instruments measuring the same construct ≥0.50 OR at least 75% of the results are in accordance with the hypotheses OR AUC ≥0.70 AND correlations with changes in related constructs are higher than with unrelated constructs
	?	Solely correlations determined with unrelated constructs
	-	Correlation with changes on instruments measuring the same construct <0.50 OR < 75% of the results are in accordance with the hypotheses OR AUC <0.70 AND correlations with changes in related constructs are higher than with unrelated constructs

+=positive rating; ?=indeterminate rating; -=negative rating; ICC=intraclass correlation coefficient; MIC=minimal important change; SDC=smallest detectable change; LOA=limits of agreement, DIF=differential item functioning

Table S2: Levels of evidence for the overall quality of the measurement properties, based on the Cochrane Back Review Group (2003)

Level	Rating	Criteria
Strong	+++ or ---	Consistent findings in multiple studies of good methodological quality
		OR in one study of excellent methodological quality
Moderate	++ or --	Consistent findings in multiple studies of fair methodological quality
		OR in one study of good methodological quality
Limited	+ or -	One study of fair methodological quality
Conflicting	±	Conflicting findings
Unknown	?	Only studies of poor methodological quality

+=positive results, ?=indeterminate results, -=negative results.

Table S3: Outcome measures used to measure hand function in children with cerebral palsy, but not included in the systematic review.

Reason excluded	**Outcome measures**
No published full-text paper on measurement properties for children with bilateral cerebral palsy, or unknown whether children with bilateral CP were included in the study population, or less than 20% of children with bilateral CP in the study sample	Assisting Hand Assessment Bayley Scales of Infant Development, 2nd edition Bayley Scales of Infant and Toddler Development, 3rd edition Box and Block Test of gross manual dexterity Caregiver Functional Use Survey Children's hand skills ability questionnaire Denver developmental screening tool, fine motor part Jebsen-Taylor Test Movement Assessment Battery for Children-2 Nine hole peg test NK dexterity board Pediatric motor activity log Posture and fine motor assessment for infants Purdue pegboard test Toddler Arm Use Test Modified Sollermann's Grip Function Test Children's assessment of participation with hands
No studies with the primary aim to evaluate measurement properties of the measure under question	Bruininks–Oseretsky Test of Motor Proficiency, 2nd edition
Primary purpose not to measure hand function; hand function assessment was part of a wider assessment of self-care, school-functioning, participation, quality of life or self-perceived competence	Activity Scale for Kids Assessment of Life Habits Assessment of Motor and Process Skills Children's Assessment of Participation and Enjoyment Minnesota Handwriting Assessment-Cerebral Palsy Paediatric Evaluation of Disability Inventory Preferences for Activities of Children School Assessment of Motor and Process Skills School Function Assessment Vineland adaptive behaviour scales, fine motor 2nd edition WeeFIM (Functional Independence Measurement Scale) Cerebral Palsy Quality of Life Child Health Questionnaire Child Caregiver Questionnaire KIDSCREEN Pediatric Outcomes Data Collection Instrument Pediatric Quality of Life Inventory Pictorial Scale of Perceived Competence and Social Acceptance for Young Children
Not available for use	Assessment of Children's Hand Skills In-hand Manipulation Test Upper Extremity Computer Adaptive Test Peg Moving Task

Paper IV

Development and validation of the Both Hands Assessment for children with bilateral cerebral palsy

ANN-KRISTIN G. ELVRUM,[1,2] BRITT-MARIE ZETHRÆUS,[3] TORSTEIN VIK,[2] LENA KRUMLINDE-SUNDHOLM[3]

[1]Department of Clinical Services, St. Olavs Hospital, Trondheim University Hospital, Norway
[2]Department of Laboratory Medicine, Children's and Women's Health, Faculty of Medicine, Norwegian University of Science and Technology, Trondheim, Norway
[3]Neuropediatric Unit, Department of Women's and Children's Health, Karolinska Institutet, Stockholm, Sweden

Correspondence to
Ann-Kristin G. Elvrum,
Department of Clinical Services, St. Olavs University Hospital, Postbox 3250 Sluppen
NO-7006 Trondheim, Norway
E-mail: ann-kristin.elvrum@ntnu.no

Word count: 3880

Abstract

Aim: To describe the development and validation of the Both Hands Assessment (BoHA) for children with bilateral cerebral palsy (BCP).

Method: The BoHA test content was developed through adaptation of the Assisting Hand Assessment. Data from 171 children with BCP (75 females; mean age: 6y6m, SD 3y2m, range 18m-12y), MACS levels I-III, were entered into Rasch measurement model analyses to evaluate internal scale validity and aspects of reliability. Additionally, the association between BoHA and MACS was investigated.

Results: Sixteen BoHA items (11 unimanual and 5 bimanual) exhibited evidence for good internal scale validity and item and person reliability when separated into two versions; one for children with asymmetric hand use (BoHA-A) and one for children with symmetric hand use (BoHA-S). The two versions were linked through anchoring of items, creating a common logit-based measure scale of bimanual performance while still allowing use of separate item difficulty hierarchies. Strong correlation was found between BoHA outcomes and MACS levels (Spearman's rho: 0.74).

Interpretation: To our knowledge, the BoHA is the first observation-based assessment of bimanual performance for children with BCP, MACS levels I-III. The BoHA has the potential to become a valuable tool in guiding treatment and measuring its effect.

What this paper adds:

- Description of a new assessment of bimanual performance (BoHA) for children with bilateral CP.
- Evidence of internal scale validity of the BoHA.
- BoHA item difficulty hierarchies that provide information useful for guiding treatment.
- Potential for using BoHA to measure effects of interventions.

Running foot: Both Hands Assessment: development and validation

Children with bilateral cerebral palsy (BCP) vary in their ability to use their hands, depending on severity and type of dominating symptom (i.e. spastic, dyskinetic or ataxic).[1,2] In addition, accompanying disturbances of sensation, perception, cognition and behavior may influence the functional use of the hands.[3] Population-based studies have indicated that about 60% of children with BCP have more than minor difficulties using their hands according to Manual Ability Classification (MACS) levels II-V.[1,2,4,5] Whereas some of these children have two relatively well-functioning hands, others have low functional abilities in both hands. Moreover, children with BCP may have an asymmetric hand function where one hand is clearly more affected than the other. This observed variation in bimanual hand use has been indicated in a population-based study by Arner and colleagues,[1] where functional ability was classified separately for each hand. However, classifications do not give detailed descriptions regarding how the hands actually are used, and such descriptions are necessary in order to describe the development of hand function and to evaluate interventions.[6] For this purpose, valid and reliable outcome measures that are responsive to change, are crucial.[7]

Most activities performed in everyday life require the use of both hands. Thus, for children with BCP impaired hand function may severely influence the children's ability to successfully perform everyday activities.[8] For children with unilateral CP (UCP) the Assisting Hand Assessment (AHA) measures how effectively the affected hand is used while performing bimanual tasks.[9] The AHA has proven to produce valid and reliable outcome measures.[10-12] In addition, the test has been found to be responsive to change, allowing monitoring of upper extremity development and evaluation of interventions in children with UCP.[11,13-15] However, little is known about how children with BCP use their hands together when handling objects. One reason for this may be the lack of outcome measures using a bimanual perspective. In a systematic review (submitted), we concluded that there are currently no available outcome measures evaluating how children with BCP use their hands together when handling objects requiring the use of both hands (bimanual performance). Thus, it would be of great interest to explore if the AHA could be modified for use with children with BCP. Since the AHA test-kit contains toys that are carefully selected to elicit spontaneous, collaborative use of the hands,[9] we hypothesized that bimanual performance could be observed using this test-kit also in children with BCP. Moreover, we hypothesized that adaptations to the AHA test items would be required to enable scoring of bimanual performance in children with different degrees of impairments in both hands.

The aim of this study was therefore to develop a new test for children with BCP, the Both Hands Assessment (BoHA), which could measure bimanual performance, as well as quantify a possible side difference between hands.

Method

The development of the BoHA involved two main steps: (1) Generation of BoHA test items, and (2) evaluation of internal scale validity and aspects of reliability of the BoHA using Rasch measurement model analyses.[16]

Participants

Eligible to participate in this study were children 18 months to 12 years old with BCP with hand function corresponding to MACS levels I-III. Children at these MACS levels can be expected to be able to handle the objects used in the BoHA test situation, although with varying degree of difficulty. Children classified to MACS levels IV-V were excluded, since they by definition have very limited, or no, ability to handle objects.[4]

The total study population comprised a convenient sample of 171 children (96 males) with BCP (mean age: 6.5 years, SD 3.2) who were recruited through pediatric habilitation units in Sweden (n=91) and Norway (n=80), see Table I. Hand dominance was established for the hand used for writing, drawing or eating. If the child did not have clear hand dominance, the dominance was characterized as “mixed”.

Informed consent was obtained from all parents and ethical approval for this study was granted by the Ethics Research Committee of Karolinska Hospital in Stockholm and by the Regional Ethical Committee (REK) for Medical Research in Mid-Norway (ref. 2012/152).

Step 1: Item generation

The content of the BoHA test items was generated from observations of bimanual play in children with BCP using the toys in the AHA test-kit to elicit spontaneous collaborative use of the hands.[9] For children aged 18 months to 5 years, the Small Kids AHA test-kit was applied using explorative play, while for 6 to 12 year old children the School Kids AHA test-kit was used with board games as the age-appropriate test session.[10,11] Furthermore, the AHA set-up for administration and video-recording was followed except for one adjustment: In the BoHA test situation, the toys were placed on both sides of the child equally often, as opposed to the AHA test situation in which most of the objects are to be placed at the child’s affected hand side.

The test developers (BMZ, LK-S and AKGE) carefully scrutinized the suitability of the original 20 AHA test items (version 5.0) and the 4-point rating scale.[12] To decide which items to include for further evaluation in the BoHA scale, the items were sorted as follows: (1) “Suitable - no changes required” when the conceptual meaning of the items and the described actions were found to be relevant and could be scored for both the dominant and the non-dominant hand; (2) “Suitable after adaptation” when the conceptual meaning of the item was found to be relevant, but the item and the categories needed some adjustment to reflect the actions observed in children with BCP; (3) “Not suitable” when the scoring categories did not fit observed performance of children with BCP. The items were evaluated independently by the test developers and discrepancies were discussed until consensus was reached.

Furthermore, new test items were generated based on object-related hand and arm actions observed in the BoHA video recordings, but that were not covered by the original AHA test items. The items and the wording of each category were developed through an iterative process of revision and observation of the children’s videotaped performances.

In addition, the clinical relevance and perceived importance of each item for evaluation of bimanual performance in children with BCP was appraised by the three test developers who independently rated each item to be: (1) essential; (2) probably essential; or (3) nonessential. Test items considered to be nonessential by all test developers were excluded.

Step 2: Evaluation of measurement properties

Rasch measurement model analyses

The Rasch measurement model for polytomous data was used to evaluate internal scale validity, including rating scale functioning, unidimensionality, targeting and aspects of reliability of the BoHA scale. Analysis was performed using the WINSTEPS computer program version 3.81.0,[17] which calculates Rasch interval logit measures from observed item and person ordinal raw scores.[16,18] The Rasch-derived interval logit measures range from negative to positive values, but can be rescaled to a more user-friendly 0 to 100 measure range.[12,18,19]

Initially, the properties of the BoHA test items were examined for acceptable rating scale functioning and for determining which derivation of the Rasch polytomous model to use for further analysis. The criteria for rating scale functioning included the following: a minimum of ten observations in each rating scale category, average category measures advancing by category, and item step calibration increasing with higher scoring categories.[20]

Secondly, unidimensionality was investigated by item and person goodness-of-fit statistics and principal components analysis (PCA) of the standardized residuals. Since misfitting infit statistics represent a larger threat to test validity than outfit, criteria for acceptable infit mean square (MnSq) values were set to infit MnSq <1.4 in combination with a standardized Z-value (Zstd) <2 for the items.[16] These criteria will identify items showing underfit to the Rasch model.[16] Items showing misfit were removed from the subsequent analyses one by one, starting with the item with the highest MnSq and Zstd values until no more than 5% of items misfit the Rasch model.[16] The criteria for unidimensionality in the item PCA were that at least 60% of the variance should be explained by the principal component (the BoHA scale), and less than 5% by any additional components.[21] The PCA of persons was used to investigate if there were contrasting groups of persons in the dataset making use of descriptive person variables such as gender, age and percentage difference in raw scores between the dominant and the non-dominant hand measured with the BoHA.[22] In addition testing for differential item functioning (DIF) was used to investigate whether the scale worked in the same way for groups with different characteristics. The DIF Contrast should be at least 0.5 logits with a statistical probability of $p<0.05$ for DIF to be noticeable.[18]

The targeting of item difficulty to person ability was then investigated by visual inspection of a person-item map, comparison of the means of item (set to 0 by default) and person measures, as well as ceiling and floor effects.[22]

In addition, the item and person reliability coefficients were investigated and the person separation ratios were calculated. The item reliability coefficient indicates the degree of replication of the item hierarchy with a different set of persons. The person reliability coefficient reflects the ability of the scale to reliably rank person's relative measure location, similar to the Cronbach's alpha. Acceptable reliability is indicated by item and person reliability coefficients ≥0.80.[23] The person separation ratio (G) was used to calculate the number of strata (ability levels) that the scale can differentiate between using the formula: (4G+1)/3.[24] A scale requires at least two distinct strata to be useful for differentiating between high and low ability.[18]

Correlation to other variables

In addition to the Rasch analyses, the association between BoHA measures and MACS levels, age, and gender were calculated by Spearman's rho and Pearson's correlations using the SPSS. The strength of the correlations was valued according to the guidelines described by Portney & Watkins,[7] where a correlation above 0.75 reflects good to excellent relationship, and a correlation between 0.50-0.75 reflects a moderate to good relationship.

Results

All children were able to complete the play session by handling all or almost all play objects in the AHA test-kit, permitting observation and scoring of bimanual performance. The play sessions took between 10-30 minutes to complete.

Item generation

Review of the 20 AHA test items revealed that five items were suitable with no changes required, nine were suitable after adaptation of the scoring criteria, while only one item (*Moves upper arm*) was regarded as not suitable (Table S1, online supporting information). In addition, the observations of hand use in the video recordings resulted in three potentially new BoHA items. These were called *Quality of arm movements, Speed of movements* and *Postural control* in sitting. Next, the appraisal of clinical relevance resulted in the exclusion of three AHA-items which showed a ceiling effect, while the original AHA items *Stabilizes by weight or support* and *Stabilizes by grasp* were merged into the modified item *Stabilizes objects*.

Thus, the item generation resulted in 18 items for the BoHA trial version: twelve unimanual and six bimanual items. The twelve unimanual items were scored separately for the dominant (D) and the non-dominant (ND) hand. The unimanual sum scores were used to determine a possible difference between hands, reported as percentage difference; the higher the number, the greater the difference between sides. For each of the six bimanual items, one common score was given for both hands. This score was added to the unimanual scores, resulting in a total of 30 data points summed up as the "Both hands sum-score", reflecting overall bimanual performance.

Evaluation of measurement properties

Rasch measurement model analyses

Initial analysis of rating scale effectiveness indicated that the BoHA trial version fulfilled the criteria for rating scale functioning using the rating scale model. The partial credit model could not be used, since not every category for each item had at least ten observations. Therefore the rating scale model was used in the further Rasch analysis.

Goodness-of-fit statistics and the PCA of items did not support unidimensionality when all 30 data points for the 171 children were included in the Rasch analysis. The PCA of persons indicated contrasting groups of persons in the dataset: In the first group, 54 out of 59 children (92%) had asymmetric hand use with a 20% or more difference between the hands; in the second group, 109 out of 110 children (99%) had a more symmetric hand use with a difference between the hands of less than 20%. In addition, the DIF analysis indicated that 87% of the items functioned differently for children with asymmetric and symmetric hand use. We therefore created two BoHA versions for the further Rasch analyses: the BoHA-S for children with symmetric hand use (<20% difference between the hands; n=116) and the BoHA-A for children with asymmetric hand use (≥20% difference between the hands; n=55). The Rasch analyses were performed in parallel for the BoHA-A and BoHA-S versions and compared in the assessment of internal scale validity of the items to ensure that the same items could be applied in both BoHA versions.

The average category measures and the step calibration measures were found to increase with the scoring category for both BoHA-S and BoHA-A versions. None of the children achieved a minimum score, whereas two (1.7%) children received a maximum score on the BoHA-S version. These two children were removed from subsequent analyses.

Goodness-of-fit statistics showed that the item *Postural control* displayed infit misfit for both BoHA versions. In addition, the item *Moves forearm* showed infit misfit for the BoHA-S version for both hands (*ND + D Moves forearm*) and for the BoHA-A version for the non-dominant hand (*ND Moves forearm*). These items were therefore removed, resulting in the acceptable level of 95% item fit (Tables II and III). Standard errors of the resulting 16 items (11 unimanual + 5 bimanual) indicated good precision of the estimates (Tables II and III). Furthermore, the PCA for the items indicated unidimensional scales with acceptable variance explained by the measure (BoHA-S: 69.1%; and BoHA-A: 73.8%). The unexplained variance in the second largest dimension was acceptable for the BoHA-S (4.5%) while it was somewhat high (6.5%) for the BoHA-A version. Although the latter improved to 4.4% when the four children with misfitting person measures were excluded, we decided from a clinical perspective to accept the slightly too high unexplained variance. The resulting goodness-of-fit statistics for the person ability measures indicated an acceptable fit to the Rasch model for 96% of the children with more symmetric hand use and 93% of the children with asymmetric hand use.

The item difficulty hierarchies of the BoHA-A and BoHA-S versions are shown in Tables II and III. The most difficult item in both versions was *Manipulates* with the non-dominant hand, while the easiest item was *Initiates use* with the dominant hand. As can be seen from the tables, the item difficulty calibrations have a larger range in the BoHA-A (4.32 to -5.78 logits) than in the BoHA-S version (4.21

to -4.56 logits). Furthermore, the person separation ratios (BoHA-A: 4.36 and BoHA-S: 5.19) and the person reliability (BoHA-A: 0.95 and BoHA-S: 0.96) indicate that the bimanual performance can be separated into 6 and 7 different ability levels, respectively.

To make it possible to report and compare the outcomes of the BoHA-A and BoHA-S versions on the same measure scale of bimanual performance, separate Rasch analyses were run where the two versions were linked using anchoring methods.[18] The item difficulty measures and item calibration structure for three items not displaying DIF (*Proceeds, ND Stabilizes objects, ND Speed of movements*) were anchored. The displacements between the anchored and unanchored logit measures were all <±0.18 logits and considerably smaller than the standard error for every item, indicating no statistical difference between the measures (Tables S2 and S3, online supporting information).[25] To make the logit measures more user-friendly they were converted to a 0-100 scale called BoHA units.[18]

The targeting of the item difficulty measures to the person ability measures expressed in BoHA units (Figure 1) were well matched and close to the average difficulty of test items (mean: 52) for children with more asymmetric hand use (mean person ability measure: 53, range: 35-75). For children with more symmetric hand use, the mean person ability measure (mean: 67, range: 41-100) was higher than the average difficulty of test items (mean: 51).

Correlation to other variables

There was a good relationship between the BoHA measures and manual abilities (MACS) (Spearman's rho: 0.74, $p \leq 0.001$). In contrast, there was low/no correlation between the BoHA measures and age (Pearson's r=0.165, $p=0.035$) and no correlation with gender (Spearman's rho=0.033, $p=0.671$).

Discussion

We found strong evidence of internal scale validity and aspects of reliability for the new outcome measure BoHA. This indicates that BoHA can measure bimanual performance in the heterogeneous group of children with BCP in the age range 18 months – 12 years and MACS levels I-III. Based on observation of bimanual play the BoHA is scored on 16 items (11 unimanual and 5 bimanual), on a 4-point rating scale. The bimanual performance is reported as an interval level logit based measure of BoHA units on a 0-100 scale. In addition, possible side differences between the hands can be quantified. The percentage difference between the hands is calculated using the ratings on the unimanual items where a higher value indicates larger asymmetry. The size of the asymmetry is used to choose the appropriate version of the outcome report sheet: The BoHA-A (asymmetric) version should be used for children with a difference ≥20% between the hands, while the BoHA-S (symmetric) version is appropriate for children with <20% difference between hands. The two versions produce comparable bimanual measures, but the item difficulty order differs between versions.

As expected, the toys in the AHA test-kit worked perfectly well to elicit bimanual play behaviour in children with BCP, while adaptations to the AHA items were necessary and revealed important differences between hand use in children with BCP and UCP. The largest conceptual difference was that items measuring decreased range or frequency of movements were not suitable for most children with BCP. In contrast, imprecise and exaggerated movements were a larger problem. In addition, slowness of movements was more pronounced in children with BCP, making bimanual performance somewhat ineffective even in children who otherwise had good abilities. Furthermore, hand role differentiation was less obvious in children with BCP. The majority of the children used both hands about equally often and the hand closest to the objects was commonly used, whether it was the dominant or the non-dominant hand. Another noteworthy difference was that the item *Coordinates* was one of the more difficult items for children with BCP, especially among children with more equilateral hand use, while this item is one of the easier items for children with UCP.[12] Besides being important for the development of the new scale, these differences also emphasize that children with UCP and BCP have differences in the use of their hands when performing bimanual activities. Therefore separate assessments are needed. The AHA produces valid measures when used for children with UCP, but for children with BCP the BoHA should be used also for those with BCP and asymmetric hand use.

The use of Rasch analyses was essential in the validation and creation of the final BoHA version, and identified two contrasting groups of children in the dataset. This necessitated the creation of the two BoHA outcome versions: one for children with clear asymmetry and one for children with more symmetric hand use. The linking of three well-fitting items using anchoring methods enabled calibration of the BoHA-A and BoHA-S versions into the same frame of reference. Although all test items are the same in both the BoHA versions, most of the items functioned differently for children with asymmetric and symmetric hand use. However, linking of three items is sufficient when the items in other respects fit the expectations of the Rasch measurement model.[18] Consequently, the BoHA can describe bimanual performance in children with BCP on the same scale regardless of degree of asymmetric hand use.

In addition to measuring bimanual performance, the BoHA provides separate item difficulty hierarchies for children with asymmetric and symmetric hand use that may be useful for treatment planning. The Rasch measurement model orders item difficulty calibrations hierarchically from easiest to hardest, as well as person ability measures from high to low ability (Tables II and III).[16] This interval scale can be used to compare a child's ability with the item difficulty levels. In this way, the items the child performs well can be identified, as well as items that are still not accomplished, but are close to the child's next ability level. The latter items may represent "just the right challenge" for the individual child, and can indicate possible targets for intervention. Moreover, the lack of correlation between the BoHA measures and the age of the children, indicates that the item difficulty hierarchies reflect steps of increasing ability rather than age-dependent development. Thus, the BoHA is a criterion-referenced test which can validly be applied for the large age range 18 months – 12 years, similar to the AHA.[11]

The ability of the BoHA-A and the BoHA-S to separate bimanual performance into 6 and 7 different ability levels, respectively, may indicate that both versions are likely to be responsive to change. Thus, the use of the BoHA may enable comparison of functional abilities in children with various types of BCP, both in studies of natural development of hand function and in the evaluation of interventions. However, further research is needed to evaluate responsiveness to change of the BoHA, as well as rater reliability (test-retest, intra- and inter-rater), and the relationship with other outcome measures

A limitation of this study was the relatively few children (n=55) with asymmetric hand use in the BoHA-A version. The sample size requirement for performing Rasch analysis is said to be sufficient around 150 participants.[26] Thus, the need to separate the children into two groups may have produced less precise estimates. However, the standard errors for both versions were generally small, indicating that the number of observations used to make the estimate was sufficient.[18] Still, further analyses including a larger number of children are desired to confirm our results. In addition further studies are required to investigate potential differences in bimanual performance for children with spastic, dyskinetic or ataxic subtypes. We included children with all subtypes in our study, but the information regarding subtype was incomplete and there were not enough subjects in each subtype to draw any conclusions.

Conclusion

The current paper reports the development of a new measure of bimanual performance, the BoHA, for children with BCP in the age range 18 months – 12 years with hand function classified to MACS levels I-III. Sixteen BoHA items exhibited strong evidence for internal scale validity and item and person reliability when separated into two versions: one for children with asymmetric hand use (BoHA-A) and one for children with more symmetric hand use (BoHA-S). By linking the BoHA measures of the two versions through anchoring of items, the measures of the respective versions are comparable, while still allowing use of separate item hierarchies. These item difficulty hierarchies indicate differences in what constitutes effective bimanual hand use for children with asymmetric or symmetric hand use, which may be useful for treatment planning. Furthermore, both BoHA versions are likely to be responsive to change.

References

1. Arner M, Eliasson AC, Nicklasson S, Sommerstein K, Hagglund G. Hand function in cerebral palsy. Report of 367 children in a population-based longitudinal health care program. *Journal of Hand Surgery (American Volume)*. 2008; **33**: 1337-47.
2. Andersen GL, Irgens LM, Haagaas I, Skranes JS, Meberg AE, Vik T. Cerebral palsy in Norway: prevalence, subtypes and severity. *European Journal of Paediatric Neurology*. 2008; **12**: 4-13.
3. Eliasson AC. Improving the use of hands in daily activities: aspects of the treatment of children with cerebral palsy. *Phys Occup Ther Pediatr*. 2005; **25**: 37-60.
4. Eliasson AC, Krumlinde-Sundholm L, Rosblad B, et al. The Manual Ability Classification System (MACS) for children with cerebral palsy: scale development and evidence of validity and reliability. *Dev Med Child Neurol*. 2006; **48**: 549-54.
5. Himmelmann K, Beckung E, Hagberg G, Uvebrant P. Gross and fine motor function and accompanying impairments in cerebral palsy. *Dev Med Child Neurol*. 2006; **48**: 417-23.
6. Krumlinde-Sundholm L. Choosing and using assessments of hand function. In: Eliasson AC, Burtner, P.A., editor. Improving hand function in children with cerebral palsy: theory, evidence and intervention. London: Mac Keith Press, 2008. p. 176-97.
7. Portney LG, Watkins MP. Foundations of clinical research: applications to practice (3rd ed). New Jersey: Pearson Education, 2009.
8. Öhrvall AM, Eliasson AC, Lowing K, Odman P, Krumlinde-Sundholm L. Self-care and mobility skills in children with cerebral palsy, related to their manual ability and gross motor function classifications. *Dev Med Child Neurol*. 2010; **52**: 1048-55.
9. Krumlinde-Sundholm L, Eliasson AC. Development of the Assisting Hand Assessment: a Rasch-built measure intended for children with unilateral upper limb impairments. *Scandinavian Journal of Occupational Therapy*. 2003; **10**: 16-26.
10. Holmefur M, Aarts P, Hoare B, Krumlinde-Sundholm L. Test-retest and alternate forms reliability of the assisting hand assessment. *J Rehabil Med*. 2009; **41**: 886-91.
11. Krumlinde-Sundholm L, Holmefur M, Kottorp A, Eliasson AC. The Assisting Hand Assessment: current evidence of validity, reliability, and responsiveness to change. *Dev Med Child Neurol*. 2007; **49**: 259-64.
12. Holmefur M, Krumlinde-Sundholm L. Psychometric properties of a revised version of the Assisting Hand Assessment (Kids-AHA 5.0). *Dev Med Child Neurol*. 2015.
13. Hoare B, Imms C, Villanueva E, Rawicki HB, Matyas T, Carey L. Intensive therapy following upper limb botulinum toxin A injection in young children with unilateral cerebral palsy: a randomized trial. *Dev Med Child Neurol*. 2013; **55**: 238-47.
14. Wallen M. Reflections on the contribution of the Assisting Hand Assessment. *Developmental Medicine & Childhood Neurology*. 2015.
15. Green D, Schertz M, Gordon AM, et al. A multi-site study of functional outcomes following a themed approach to hand-arm bimanual intensive therapy for children with hemiplegia. *Dev Med Child Neurol*. 2013; **55**: 527-33.
16. Bond TG, Fox CM. Applying the Rasch model: fundamental measurement in the human sciences. Second edition ed. New York: Taylor & Francis Group, LLC, 2007.
17. Linacre JM. Winsteps ®. 3.81.0 ed. Beaverton, Oregon: Winsteps.com. Retrieved January 1, 2014. Available from http://www.winsteps.com; 2014.
18. Linacre JM. Winsteps ® Rasch measurement computer program User's Guide. Beaverton, Oregon: Winsteps.com2014. Report No.: 0-941938-03-4.
19. Krumlinde-Sundholm L. Reporting outcomes of the Assisting Hand Assessment: what scale should be used? *Dev Med Child Neurol*. 2012; **54**: 807-8.
20. Linacre JM. Optimizing rating scale category effectiveness. *Journal of Applied Measurement*. 2002; **3**: 85-106.
21. Smith EV, Jr. Detecting and evaluating the impact of multidimensionality using item fit statistics and principal component analysis of residuals. *Journal of Applied Measurement*. 2002; **3**: 205-31.

22. Tennant A, Conaghan PG. The Rasch measurement model in rheumatology: what is it and why use it? When should it be applied, and what should one look for in a Rasch paper? *Arthritis Rheum*. 2007; **57**: 1358-62.
23. Fisher Jr W. Rating Scale Instrument Quality Criteria. *Rasch Measurement Transactions*. 2007; **21**: 1095.
24. Wright B MG. Rating scale analysis. Chicago: Mesa Press, 1982.
25. O'Neill T, Peabody, M, Tan, RJB, Du, Y How much item drift is too much? *Rasch Measurement Transactions*. 2013; **27**: 1423-4.
26. Linacre JM. Sample Size and Item Calibration Stability. *Rasch Measurement Transactions*. 1994; **7**: 328.

Table I: Demographic characteristics of the included participants (n=171) with bilateral cerebral palsy (BCP)

	Total n (%)	Symmetry n (%)	Asymmetry n (%)
BoHA assessments	171 (100)	116 (68)	55 (32)
Age			
18-60 months	66 (39)	37 (32)	29 (53)
61 months – 12 years	105 (61)	79 (68)	26 (47)
Gender			
Males	96 (56)	65 (56)	31 (56)
Females	75 (44)	51 (44)	24 (44)
Hand dominance			
Right hand	109 (64)	79 (68)	30 (55)
Left hand	43 (25)	18 (16)	25 (45)
Mixed	19 (11)	19 (16)	0
MACS*			
I	53 (32)	48 (43)	5 (10)
II	56 (34)	38 (34)	18 (34)
III	55 (34)	26 (23)	29 (56)

BoHA= Both Hands Assessment; n=number; MACS=Manual Ability Classification System; *=7 missing MACS classifications

Table II: Item difficulty order and item fit statistics for the Both Hands Assessment Symmetry (BoHA-S) version consisting of 11 unimanual items scored separately for the dominant (D) and the non-dominant (ND) hand and 5 bimanual (B) items scored with one common score for both hands.

	Measure		Infit		Outfit		PMC
Item	Logits	SE	MnSq	Zstd	Mnsq	Zstd	
ND Manipulates	4.20	0.22	1.29	1.9	1.46	2.1	0.62
ND Grip force regulation	2.54	0.21	0.74	-2.0	0.65	-2.1	0.83
D Manipulates	2.27	0.21	1.15	1.1	1.07	0.4	0.73
D Grip force regulation	1.96	0.21	1.02	0.2	0.93	-0.4	0.77
B Proceeds	1.65	0.21	1.30	2.1	1.40	2.1	0.79
B Flow in bimanual performance	1.47	0.21	0.84	-1.2	0.84	-0.9	0.89
B Coordinates	1.29	0.21	0.85	-1.1	0.86	-0.8	0.86
ND Stabilizes objects	0.90	0.21	1.04	0.4	1.00	0.1	0.78
ND Moves fingers	0.81	0.21	0.80	-1.6	0.79	-1.1	0.82
ND Varies type of grasp	0.59	0.21	0.78	-1.7	0.76	-1.2	0.84
B Readjusts grasp	0.55	0.21	0.75	-2.1	0.65	-2.0	0.86
D Moves fingers	0.37	0.21	0.74	-2.2	0.78	-1.0	0.83
B Orients objects	0.10	0.21	0.95	-0.3	0.89	-0.4	0.76
ND Quality of arm movements	-0.04	0.21	1.12	0.9	1.05	0.3	0.78
ND Grasps	-0.08	0.21	0.79	-1.6	0.65	-1.5	0.85
D Stabilizes objects	-0.13	0.21	1.26	1.9	1.22	0.9	0.73
D Varies type of grasp	-0.27	0.22	1.1	0.8	1.05	0.3	0.76
D Quality of arm movements	-0.40	0.22	1.12	0.9	1.03	0.2	0.76
ND Speed of movements	-0.50	0.22	1.08	0.7	0.98	0.0	0.77
D Speed of movements	-0.59	0.22	1.07	0.5	0.97	0.0	0.76
ND Releases	-1.34	0.23	0.69	-2.4	0.58	-1.1	0.81
D Grasps	-1.44	0.23	0.81	-1.4	0.64	-0.9	0.79
ND Reaches	-1.71	0.24	1.33	2.0	1.15	0.5	0.68
D Releases	-2.06	0.24	0.49	-4.1	0.41	-1.6	0.81
ND Initiates	-2.48	0.25	1.35	2.0	1.43	0.9	0.61
D Reaches	-3.09	0.27	1.46	2.4	1.19	0.5	0.60
D Initiates	-4.56	0.33	1.37	1.7	0.66	-0.3	0.52
Mean	0.00	0.22	1.01	-0.1	0.93	-0.3	
Standard deviation (SD)	1.83	0.02	0.25	1.7	0.26	1.1	

SE=standard error; MnSq=Mean Square; Zstd=Z score; PMC=point measure correlation

Table III: Item difficulty order and item fit statistics for the Both Hands Assessment Asymmetry (BoHA-A) version consisting of 11 unimanual items scored separately for the dominant (D) and the non-dominant (ND) hand and 5 bimanual (B) items scored with one common score for both hands.

	Measure		Infit		Outfit		PMC
Item	Logits	SE	MnSq	Zstd	Mnsq	Zstd	
ND Manipulates	4.32	0.31	0.93	-0.3	0.84	-0.6	0.59
ND Varies type of grasp	3.86	0.30	0.90	-0.5	0.86	-0.6	0.71
B Flow in bimanual performance	3.23	0.30	0.73	-1.5	0.64	-1.8	0.87
ND Grip force regulation	3.06	0.30	0.98	0.0	0.97	-0.1	0.66
B Readjusts grasp	2.88	0.29	1.01	0.1	1.03	0.2	0.72
ND Reaches	2.71	0.29	1.54	2.5	1.63	2.5	0.53
ND Grasps	2.54	0.29	1.06	0.4	1.03	0.2	0.48
B Coordinates	1.94	0.29	0.95	-0.2	0.91	-0.4	0.85
ND Moves fingers	1.77	0.29	0.60	-2.3	0.56	-2.4	0.62
ND Releases	1.61	0.29	0.82	-0.9	0.81	-0.9	0.69
B Proceeds	1.52	0.29	1.20	1.0	1.19	0.9	0.86
ND Quality of arm movements	1.52	0.29	0.51	-2.9	0.44	-3.2	0.73
ND Stabilizes objects	1.36	0.29	0.82	-0.9	0.77	-1.1	0.71
B Orients objects	1.11	0.29	1.24	1.2	1.30	1.4	0.75
ND Initiates	0.95	0.29	0.76	-1.3	0.81	-0.8	0.69
ND Speed of movements	0.07	0.28	0.87	-0.6	0.85	-0.7	0.85
D Manipulates	-0.84	0.27	1.19	1.0	1.15	0.8	0.73
D Grip force regulation	-1.44	0.27	1.18	1.0	1.21	1.0	0.59
D Varies type of grasp	-2.10	0.27	1.05	0.3	1.03	0.2	0.72
D Moves fingers	-2.40	0.27	0.99	0.0	0.96	-0.1	0.63
D Speed of movements	-2.93	0.28	0.97	-0.1	0.94	-0.1	0.62
D Stabilizes objects	-3.01	0.28	1.32	1.7	1.32	1.1	0.63
D Grasps	-3.48	0.29	0.91	-0.4	1.02	0.2	0.70
D Quality of arm movements	-3.57	0.29	1.17	0.9	1.14	0.5	0.62
D Releases	-4.26	0.30	0.77	-1.2	0.72	-0.5	0.63
D Reaches	-4.64	0.32	1.35	1.6	1.11	0.4	0.53
D Initiates use	-5.78	0.37	1.05	0.3	0.82	0.0	0.47
Mean	0.00	0.29	1.00	0.0	0.97	-0.1	
Standard deviation (SD)	2.87	0.02	0.23	1.2	0.25	1.1	

SE=standard error; MnSq=Mean Square; Zstd=Z score; PMC=point measure correlation

Figure legends

Figure 1: Variable map illustrating the targeting of item difficulty to person ability for children with bilateral cerebral palsy assessed with the Both Hands Assessment (BoHA). The targeting for children with more symmetrical hand use is illustrated on the left side of the figure (BoHA-Symmetry), while the targeting of children with asymmetrical hand use is illustrated on the right side of the figure (BoHA-Asymmetry). The more able persons and more difficult items are at the top of the map, and lower performing persons and easier items are at the bottom. X represents one person or one item. The item measures are shown on their threshold values between a score of 1 or 2, 2 or 3, and 3 or 4, respectively. Measures are given in BoHA units.

Acknowledgement

This research was funded by Sunnerdahls Handikappfond, Norrbacka-Eugenisastiftelsen, Stiftelsen Frimurare Barnhuset i Stocholm and the Liaison Committee made up of representatives from the Central Norway Regional Health Authority (RHA) and the Norwegian University of Science and Technology (NTNU).

The second and the last author are both involved in the company Handfast AB (Inc.) which supplies training courses and test material for the Assisting Hand Assessment. The company has not been involved in or funded any part of the research.

Supporting information

Additional supporting information may be found in the online version of this article:

Table S1: Item generation for the Both Hands Assessment (BoHA).
Table S2: Item difficulty order and item fit statistics for the anchored Both Hands Assessment Symmetry (BoHA-S) subscale.
Table S3: Item difficulty order and item fit statistics for the anchored Both Hands Assessment Asymmetry (BoHA-A) subscale.

```
               BoHA-Symmetry                                BoHA-Asymmetry

    Person |         ITEM SCORES       BoHA-units           ITEM SCORES       | Person
           | 1-2   | 2-3   | 3-4                    1-2   | 2-3   | 3-4     |
        XX |       |       |           100           |       |       |
          ------------------------------------------------------------------
           |       |       |                         |       |       |
           |       |       |                         |       |       |
           |       |       |                         |       |       |
           |       |       |                         |       |       |
      XXXX |       |       |                         |       |       |
           |       |       |                         |       |       |
           |       |       |            90           |       |       |
          ------------------------------------------------------------------
         X |       |       |                         |       |       |
           |       |       | X                       |       | X     |
         X |       |       |                         |       | X     |
           |       |       |                         |       |       |
     XXXXX |       |       |                         |       | XX    |
        XX |       |       | X                       |       | XX    |
   XXXXXXX |       |       | X          80           |       | X     |
          ------------------------------------------------------------------
         X |       |       | X                       |       |       |
      XXXX |       |       | XXX                     |       | XXXXX |
        XX |       |       |                         |       | X     | X
         X |       |       | XX                      |       | XX    |
XXXXXXXXXX |       |       | XXX                     |       |       |
       XXX |       |       | X                       |       |       | X
  XXXXXXXX |       |       | XXXX       70           | X     | X     | X
          ------------------------------------------------------------------
           |       | X     | XXX                     | X     |       | X
   XXXXXXX |       |       |                         |       | X     |
      XXXX |       |       |                         | XX    |       | X
   XXXXXXX |       |       | XX                      | XX    | X     | XX
        XX |       | X     | X                       | X     |       | XX
       XXX |       | X     | X                       |       | X     | XX
    XXXXXX |       | X     | X          60           | XXXXX | X     | X
          ------------------------------------------------------------------
    XXXXXX |       | XXX   |                         | X     | X     | X
      XXXX |       |       | X                       | XX    | X     | XXXXX
     XXXXX |       | XX    |                         |       | XX    | XXXXXX
           |       | XXX   |                         |       |       | XXX
           |       | X     |                   X     | X     | X     | XXXXXX
    XXXXXX |       |       |                   X     |       |       | XX
      XXXX |       | XXXX  | X          50           | X     | X     | XX
          ------------------------------------------------------------------
        XX |       | XXX   |                   XX    |       |       | XXX
       XXX | X     |       |                   XX    | X     |       | XXXXXX
        XX |       |       |                   X     |       | X     | X
         X |       | XX    |                         |       |       | XXX
         X |       | X     |                         | X     |       |
         X | X     | X     |                   XXXXX | X     |       | XX
         X | X     | X     |            40     X     | X     |       | XX
          ------------------------------------------------------------------
           | X     |       |                   XX    | X     |       |
           | XXX   | X     |                         | XX    |       |
           |       |       |                         |       |       |
           | XX    |       |                   X     |       |       | X
           | XXX   |       |                         | X     |       |
           | X     | X     |                   X     |       |       |
           | XXXX  |       |            30           | X     |       |
          ------------------------------------------------------------------
           | XXX   |       |                   X     |       |       |
           |       |       |                         | X     |       |
           |       |       |                   X     |       |       |
           | XX    |       |                   X     |       |       |
           | X     |       |                   X     |       |       |
           | X     |       |                   X     |       |       |
           | X     |       |            20     XX    |       |       |
          ------------------------------------------------------------------
           |       |       |                         |       |       |
           | X     |       |                   X     |       |       |
           |       |       |                         |       |       |
           |       |       |                   X     |       |       |
           |       |       |                         |       |       |
           | X     |       |                         |       |       |
           |       |       |            10     X     |       |       |
          ------------------------------------------------------------------
           |       |       |                         |       |       |
           |       |       |                         |       |       |
           |       |       |                         |       |       |
           |       |       |                         |       |       |
           |       |       |             0           |       |       |
 Less able |       |       |       Easier item       |       |       | Less able
```

Figure 1: Variable map illustrating the targeting of item difficulty to person ability for children with bilateral cerebral palsy assessed with the Both Hands Assessment (BoHA). The targeting for children with more symmetrical hand use is illustrated on the left side of the figure (BoHA-Symmetry), while the targeting of children with asymmetrical hand use is illustrated on the right side of the figure (BoHA-Asymmetry). The more able persons and more difficult items are at the top of the map, and lower performing persons and easier items are at the bottom. X represents one person or one item. The item measures are shown on their threshold values between a score of 1 or 2, 2 or 3, and 3 or 4, respectively. Measures are given in BoHA units.

Table S1: Item generation for the Both Hands Assessment (BoHA) based on evaluation of the suitability of the original Assisting Hand Assessment (AHA) test items, as well as generation of potential new test items, and appraisal of clinical relevance.

Suitable AHA items	Suitable AHA items after adaptation	Not suitable AHA items	Potentially new items	Items of less clinical relevance
• Moves forearm • Varies type of grasp • Grip force regulation • Readjusts grasp • Orients objects	• Initiate • Reaches • Grasps • Releases • Moves fingers • Manipulates • Coordinates • Proceeds • Flow in bimanual task performance • Stabilizes objects (Stabilizes by grasp & Stabilizes by weight or support)	• Moves upper arm	• Postural control • Quality of arm movements • Speed of movements	• Amount of use • Holds • Chooses assessed hand when closer to objects

Table S2: Item difficulty order and item fit statistics for the anchored Both Hands Assessment Symmetry (BoHA-S) subscale consisting of 11 unimanual items scored separately for the dominant (D) and the non-dominant (ND) hand and 5 bimanual items scored with one common score for both hands (B). Three items are anchored (A).

Item	Logits	SE	Infit		Outfit		Displace
			MnSq	Zstd	Mnsq	Zstd	
ND Manipulates	3.21	0.18	1.03	0.3	1.05	0.4	0.00
ND Grip force regulation	2.08	0.18	0.56	-4.1	0.53	-4.1	0.00
D Manipulates	1.89	0.18	0.87	-1.0	0.85	-1.1	0.00
D Grip force regulation	1.66	0.18	0.77	-1.9	0.73	-2.1	0.00
B Proceeds	1.33A	0.18	0.94	-0.4	1.05	0.4	0.11
B Flow in bimanual performance	1.30	0.18	0.60	-3.5	0.57	-3.5	0.00
B Coordinates	1.17	0.18	0.63	-3.2	0.59	-3.2	0.00
ND Stabilizes objects	0.91A	0.18	0.80	-1.6	0.75	-1.7	-0.04
ND Moves fingers	0.80	0.18	0.61	-3.4	0.67	-2.3	0.00
ND Varies type of grasp	0.63	0.19	0.61	-3.4	0.60	-2.7	0.00
B Readjusts grasp	0.60	0.19	0.59	-3.7	0.53	-3.3	0.00
D Moves fingers	0.46	0.19	0.58	-3.7	0.65	-2.2	0.00
B Orients objects	0.25	0.19	0.75	-2.1	0.73	-1.5	0.00
ND Quality of arm movements	0.14	0.19	0.88	-0.9	0.78	-1.2	0.00
ND Grasps	0.11	0.19	0.64	-3.1	0.56	-2.6	0.00
D Stabilizes objects	0.07	0.19	0.99	0.0	0.91	-0.4	0.00
D Varies type of grasp	-0.04	0.19	0.86	-1.0	0.83	-0.8	0.00
D Quality of arm movements	-0.15	0.19	0.88	-0.9	0.77	-1.0	0.00
ND Speed of movements	-0.16A	0.19	0.85	-1.1	0.76	-1.1	-0.07
D Speed of movements	-0.30	0.19	0.84	-1.2	0.75	-1.1	0.00
ND Releases	-0.89	0.20	0.58	-3.5	0.51	-1.8	0.00
D Grasps	-0.97	0.21	0.66	-2.7	0.56	-1.5	0.00
ND Reaches	-1.19	0.21	1.03	0.2	0.88	-0.2	0.00
D Releases	-1.46	0.22	0.43	-4.9	0.37	-2.0	0.00
ND Initiates	-1.80	0.22	1.01	0.1	0.98	0.1	0.00
D Reaches	-2.28	0.24	1.12	0.8	0.87	-0.1	0.00
D Initiates	-3.44	0.29	1.05	0.3	0.53	-0.6	0.00
Mean	0.15	0.20	0.78	-1.8	0.72	-1.5	
Standard deviation (SD)	1.41	0.02	0.18	1.6	0.17	1.2	

SE=standard error; MnSq=Mean Square; Zstd=Z score; PMC=point measure correlation

Table S3: Item difficulty order and item fit statistics for the anchored Both Hands Assessment Asymmetry (BoHA-A) subscale consisting of 11 unimanual items scored separately for the dominant (D) and the non-dominant (ND) hand and 5 bimanual items scored with one common score for both hands (B). Three items are anchored (A).

Item	Logits	SE	Infit		Outfit		Displace
			MnSq	Zstd	Mnsq	Zstd	
ND Manipulates	3.49	0.29	0.82	-1.0	0.78	-1.0	0.00
ND Varies type of grasp	3.09	0.28	0.77	-1.3	0.73	-1.4	0.00
B Flow in bimanual performance	2.56	0.27	0.60	-2.5	0.55	-2.7	0.00
ND Grip force regulation	2.41	0.27	0.81	-1.0	0.81	-1.0	0.00
B Readjusts grasp	2.27	0.27	0.84	-0.8	0.83	-0.8	0.00
ND Reaches	2.12	0.27	1.28	1.5	1.30	1.5	0.00
ND Grasps	1.98	0.27	0.90	-0.5	0.88	-0.6	0.00
B Coordinates	1.48	0.26	0.74	-1.4	0.72	-1.5	0.00
ND Moves fingers	1.35	0.26	0.53	-2.9	0.51	-3.0	0.00
B Proceeds	1.33A	0.26	0.95	-0.2	0.94	-0.3	-0.18
ND Releases	1.21	0.26	0.68	-1.8	0.67	-1.8	0.00
ND Quality of arm movements	1.14	0.26	0.45	-3.6	0.42	-3.8	0.00
ND Stabilizes objects	0.91A	0.26	0.68	-1.8	0.66	-1.9	0.09
B Orients objects	0.80	0.26	1.00	0.1	1.02	0.1	0.00
ND Initiates	0.67	0.26	0.64	-2.1	0.65	-2.0	0.00
ND Speed of movements	-0.16A	0.26	0.74	-1.5	0.72	-1.5	0.10
D Manipulates	-0.84	0.25	1.02	0.2	1.00	0.1	0.00
D Grip force regulation	-1.36	0.25	1.01	0.1	1.02	0.2	0.00
D Varies type of grasp	-1.95	0.26	0.93	-0.3	0.90	-0.4	0.00
D Moves fingers	-2.21	0.26	0.86	-0.8	0.83	-0.8	0.00
D Speed of movements	-2.69	0.26	0.86	-0.7	0.83	-0.7	0.00
D Stabilizes objects	-2.76	0.26	1.20	1.1	1.17	0.8	0.00
D Grasps	-3.19	0.27	0.84	-0.8	0.88	-0.4	0.00
D Quality of arm movements	-3.26	0.27	1.06	0.4	1.01	0.1	0.00
D Releases	-3.89	0.29	0.68	-1.8	0.65	-0.9	0.00
D Reaches	-4.23	0.30	1.21	1.0	1.01	0.2	0.00
D Initiates	-5.28	0.35	0.97	-0.1	0.74	-0.2	0.00
Mean	-0.19	0.27	0.86	-0.8	0.82	-0.9	
Standard deviation (SD)	2.47	0.02	0.20	1.2	0.20	1.2	

SE=standard error; MnSq=Mean Square; Zstd=Z score; PMC=point measure correlation