UNIVERSITÀ DEGLI STUDI FIRENZE

DOTTORATO TOSCANO DI NEUROSCIENZE

CICLO XXX

COORDINATORE Prof. Corradetti

Quantitative measures of spontaneous and intervention-induced perceptual-motor development in children with neurodevelopmental disorders

Settore Scientifico Disciplinare MED/39

Dottoranda
Elena Beani

Tutor
Prof. Giovanni Cioni

Co-Tutor
Dr.ssa Giuseppina Sgandurra

Coordinatore
Prof. Renato Corradetti

Anni 2014/2018

Ai bambini, tutti

Ai bambini che stanno combattendo contro mali più grandi di loro

Alle mie nipotine "non di sangue, ma di cuore", che mi fanno commuovere

Ai miei campioni, che mi insegnano più di quanto io riesca ad insegnare a loro

A tutti i bimbi che mi vedono diventare grande, ma sanno che rimango una di loro!

Index

1. Introduction
    1.1. Neurodevelopmental disorders……..……..1
        1.1.1.Definition…….……...1
        1.1.2.Early detection and diagnosis of neurodevelopmental disorders……..3
        1.1.3.Early intervention and evaluation of outcome……....5
    1.2. Outcome measures…….....7
        1.2.1.Outcome measures: general overview……....7
        1.2.2.Measurement properties…….……..10
        1.2.3.ICF framework…….12
        1.2.4.Clinical vs technological outcome measures……14
            1.2.4.1. Visual assessment…….14
            1.2.4.2. Assessment of upper limb movements…….16
            1.2.4.3. General considerations…….22
    1.3. Purposes of the thesis……..….25
2. CareToy C: A new system for quantitative evaluation of infant gaze capabilities in a wide visual field…….28
    2.1. Introduction …….29
    2.2. Materials and methods……..31
    2.3. Results ……41
    2.4. Discussion……44
    2.5. Final remarks……....47
3. CareToy H: Quantitative assessment and intervention approaches for infants at risk for neurodevelopmental disorders…….49
    3.1. Introduction……..50
    3.2. Sensorized toys for measuring manipulation capabilities of infants at home…….54
        3.2.1.Materials and methods……..54

3.2.2.Results and discussion……..58
3.3. CareToy: Stimulation and assessment of preterm infant's activity using a novel sensorized system……...60
3.3.1.Materials and methods…….….60
3.3.2.Results …….71
3.3.3.Discussion…….75
3.4. A randomized clinical trial in preterm infants on the effects of a home-based early intervention with the 'CareToy system'. …….79
3.4.1.Materials and methods……..80
3.4.2.Results …….85
3.4.3.Discussion…….88
3.5. Sensorized toys for measuring the effect of home-based therapy on manipulation capabilities of infants…….93
3.5.1.Materials and methods…….93
3.5.2.Results ……...100
3.5.3.Discussion……...103
3.6. CareToy Early Intervention in infants with Down Syndrome: a feasibility study ……105
3.6.1.Materials and methods……105
3.6.2.Results ……...110
3.6.3.Discussion……...116
3.7. Final remarks……..119
4. Quantitative assessment of upper limb use in children with typical development and with neurodevelopmental disorders by inertial sensors ……122
4.1. Introduction……123
4.2. Assessment of upper limb use in children with typical development and with neurodevelopmental disorders by inertial sensors: a systematic review……...126
4.2.1.Methods……..126
4.2.2.Results ……...130

4.2.3.References of the review……164
4.3. Asymmetry of Upper Limb activity in children with unilateral cerebral palsy: validation of a Triaxial Accelerometer Approach……168
4.3.1.Materials and methods……169
4.3.2.Results ……...174
4.3.3.Discussion……...179
4.4. Final remarks……..182
5. Action-Observation Therapy (AOT) and Information and Communications Technologies (ICT) for Home Rehabilitation of Children with Hemiplegia……..184
5.1. Action observation for upper and lower limb rehabilitation in hemiplegia: a systematic review…….185
5.1.1. Methods…….185
5.1.2.Results……189
5.1.3.Discussion……...216
5.1.4.References of the review……232
5.2. Tele-UPCAT: Study Protocol of a Randomized Controlled Trial of a Home-Based Tele-monitored UPper Limb Children Action Observation Training for Participants with Unilateral Cerebral Palsy……...……238
5.2.1.Methods……..238
5.3. A pilot study on a sample group of children with unilateral cerebral palsy using tele-UPCAT…….254
5.3.1.Materials and methods……255
5.3.2.Results……266
5.3.3.Discussion ……..272
5.4. Acceptability and usability of tele-UPCAT system for Action observation training in children with Unilateral Cerebral Palsy…276
5.4.1.Methods……..276
5.4.2.Results ……...281

5.4.3.Discussion ……..285
5.5. Final remarks……..288
6. Summary and conclusions……..292
7. Acknowledgments …….296
8. References……..300

CHAPTER 1
INTRODUCTION

1.1 Neurodevelopmental disorders

1.1.1 Definition

Neurodevelopmental disorders are impairments of brain growth and development affecting several brain functions, and include cognitive, motor, language, learning, and behavioural disorders due to many causes – genetic, lesional, and environmental (Cioni et al, 2016).

Neurodevelopmental disorders encompass a group of clinical heterogeneous conditions with onset in the developmental period. These disorders typically manifest early in development and are characterized by developmental deficits that produce lifetime impairments of personal, social, and occupational functioning. To date, the etiological bases of the majority of these conditions are still unknown, though a great body of data supports their polygenic and multifactorial etiology (De Felice et al, 2015), although premature birth represents one of the risk factors. The incidence of preterm births has increased and survival rates in very preterm infants have improved over the past two decades (Doyle, 2004). However, a significant number of these infants, especially if born with a very low birth weight (VLBW), show a developmental disorder. Follow-up studies indicate that about 50% show cognitive, motor, or behavioural problems in childhood (Spittle et al, 2012), resulting in a lower high school graduation rate compared with infants born at term with normal birth weight (Hack et al, 2002); a percentage up to 15% of preterms are diagnosed with Cerebral Palsy. Cerebral Palsy is 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 (Blair et al, 2006).

Cerebral palsy is the most common cause of childhood chronic physical disability in Europe and in other industrialised societies (Krägeloh-Mann and Cans). The incident rate is between 2 and 3 per 1000 live births and increases to 40–100 per 1000 live births among very premature and very low birth-weight babies which represent the population with highest rate of neurodevelopmental disorders (Stavsky et al, 2017). Unilateral Cerebral Palsy (UCP) (i.e. a motor impairment impacting one side of the body) constitutes the most frequent form of Cerebral Palsy, about 30-40% of all affected children (Novak et al, 2017). Recent estimations of incidence and prevalence of CP have shown a significant increase in UCP in Europe over the last years (Andersen et al, 2011). The Upper Limb (UL) of children with UCP is generally more involved than the lower limb and the consequent disability affects their participation, quality of life, independence at home, school and community.

1.1.2 Early detection ad diagnosis of neurodevelopmental disorders

Knowing the most common neurodevelopmental disorders and their consequences highlights the importance of the early detection of these pathologies. In addition, it has been demonstrated how many of these conditions respond well to interventions in early childhood, when brain plasticity is at its greatest (Johnston et al, 2009) and developmental trajectories can be modified with maximal benefit into adulthood (Heckman, 2008).

In fact, there is a crucial aspect to be considered that is the neural plasticity which, in general terms, refers to functional and structural changes in the brain, which are advantageous to offset or improve functions; the term denotes several capacities including the ability to adapt to changes in the environment and to store information in memory associated with learning (Johnston, 2004). There is abundant evidence that the structure of certain brain circuits can change in response to environmental stimuli (Chen et al, 2002).

The brain plasticity, consisting in the modifications to the central nervous system in response to environmental stimulation, allow us to learn new skills, remember new information, and recover from brain injury (Pascual-Leone et al, 2005); anyway vulnerability to adverse environmental factors represents the drawback of brain plasticity.

Based on this, one goal of child health development screening programs is to identify, as early as possible, infants at risk for future motor, visual, cognitive problems or more in general, developmental disabilities, in order to make a rehabilitation intervention possible and incisive (Morante et al, 1982).

Infants at high risk for neurodevelopmental disorders can be identified early, i.e. in the first weeks or months of life, through careful clinical evaluation (i.e. screening programs, developmental tests, neurological examination, observation of spontaneous movement patterns) combined with specific technical tools such as neuroimaging (cranial ultrasounds, brain magnetic

resonance imaging (MRI)), neurophysiological tests (e.g. electroencephalography, evoked potentials), and genetic tests (karyotype, comparative genomic hybridization-microarray). The application of evidence-based recommendations or decision-making processes, which combine the use of clinical and technical tools at a proper point during development, is crucial for early detection of infants at risk for neurodevelopmental disorders by clinicians.

An example of early identification of infants at risk for cerebral palsy (CP), one of the most common neurodevelopmental disorders (Vohr et al, 2005), is the combined use of Prechtl's General Movement Assessment and brain MRI. For the diagnosis of Cerebal Palsy, ages from 12 to 24 months were historically regarded as the latent or silent period where it could not be identified accurately. Experts now consider the silent period as outdated because cerebral palsy or "high risk of cerebral palsy" can be accurately predicted before age 6 months' corrected age. The 3 tools with best predictive validity for detecting cerebral palsy before 5months' corrected age are: neonatalmagnetic resonance imaging (MRI) (86%-89% sensitivity); Prechtl Qualitative Assessment of General Movements (GMs) (98% sensitivity), and the Hammersmith Infant Neurological Examination (HINE) (90% sensitivity). After 5 months' corrected age, the most predictive tools for detecting risk are MRI (86%-89% sensitivity) (where safe and feasible), the HINE (90% sensitivity), and the Developmental Assessment of Young Children (83% C index). High-quality evidence also indicates that a trajectory of abnormal GMs or HINE scores, in combination with abnormal MRI, producing congruent findings, is even more accurate than individual clinical assessments in isolation (Bosanquer et al, 2013). To make an early clinical diagnosis before 6 months' corrected age, a combination of assessments with strong predictive validity coupled with clinical reasoning is recommended. A highly experienced clinical team should ideally conduct and interpret the standardized assessments.

1.1.3 Early intervention and evaluation of outcome

Early detection and diagnosis of developmental disabilities allows to begin an Early intervention during the critical period of brain plasticity, improving neurodevelopmental outcomes and enhancing brain plasticity (Spittle et al, 2009). As stated by the World Health Organization, identification of the infant at risk for a neurodevelopmental disorder is a crucial starting point to establish a close relationship between parents and health care providers and to provide early intervention (WHO, 2011). The goal of early intervention is to prevent or minimize motor, cognitive, emotional impairments in young children disadvantaged by biological or environmental risk factors (Guralnick, 2011). Evidence suggests that brain development should be considered the result of the complex interaction between genes, social and physical environment thanks to the modulation of gene transcription and expression (epigenetic mechanisms) (Sale et al, 2014; Berardi et al, 2015). In this framework, environment, social relationship, and parents have a key role in early intervention. Early developmental interventions are used in clinical setting with the aim of improving the overall outcome for infants at risk or with neurodevelopmental disorders (Lundqvist-Persson, 2012). Early Intervention (EI) means intervening as soon as possible to tackle problems that have already emerged due to perinatal or congenital brain disorders. It is carried out in a critical period of development, in a time window during which specific functions develop very rapidly, when initial signs of atypical development are present but before they become overt. The target of early intervention programs is to strengthen brain plasticity, higher in the first years of life, and to allow a better functional outcome. Based on our current neurobiological understanding, early intervention programs must be implemented very early in life and should be intensive, repetitive, incrementally challenging and individualized. They should include also goal-directed components and exercises where meaningful goals are provided, to give opportunities for

problem solving and to indirectly drive the movements required to successfully meet task demands (Siegert and Taylor, 2004; Löwing et al, 2010). The therapeutic approach should also involve repeated practice of tasks, which the infant wants to do (Riethmuller et al, 2009). There is also accumulating evidence of the importance of an enriched environment (EE) on infant's neurodevelopment (Sale et al, 2009).

Home enriched environment should be organized to encourage the infant to perform specific tasks, tailored on the developmental needs of the individual infant, in an environment where the parent is actively and positively engaged with the child, to facilitate and promote learning. The home environment should include safe toys appropriate to the infant's ability level but posing a learning challenge and family interactions. In this context, as proposed by Morgan et al. (2013), enriched environment intervention could be defined as a set of modifications aimed to enrich motor, cognitive, sensory, or social aspects of the infant's environment with the purpose of promoting learning.

The early intervention is one of the basis of the treatment of neurodevelopmental disorders, but rehabilitation must continue also later, when the disorder is structured, but there are clinical signs and/or comorbidity and it is still possible to improve the clinical outcome. Main guidelines recommended in fact to follow children with neurodevelopmental disorders (as cerebral palsy, learning disabilities, attention deficits/hyperactivity disorder etc.) during all their development, by providing the more suitable and useful intervention related to their age and developmental needs. There are several models of intervention for the different disabilities and knowing the effects of each model and observe the longitudinal story of patients, help in establish which intervention is better and add knowledge for improve future rehabilitative approach. Together with this it raises up the need to accurately evaluate the efficacy of the intervention, indeed, it is crucial to have clinical and instrumental tools which allow to measure the outcome, and for this reason research is constantly focused on the exploration and the study of tools which can be used for these purposes.

1.2 Outcome measures

1.2.1 Outcome measures: general overview

Evidence based medicine (EBM) is a systematic approach to clinical problem solving which allows the integration of the best available research evidence with clinical expertise and patient values and it is the basis of the research and strictly correlates with the clinical practice. EBM approach is made up of five steps which are: i) ask a specific clinical question, ii) find the best evidence to answer the question, iii) critically appraise the evidence for its validity and usefulness, iv) integrate appraisal results with clinical expertise and patient values and v) evaluate the outcomes.

The development of effective outcome measures has become a major issue in the field of rehabilitation and it has contributed to a better understanding of how results could be linked to specific treatment elements. Measuring an effect is a complex phenomenon, whose causal factors cannot be acquired by a single evaluation tool. In general, it would be more appropriate to analyze each different result with the most appropriate measuring instrument.

The assessment of the outcome allows clinicians to: i) identifying and characterizing signs and symptoms, structural and functional limitations due to the clinical situation; ii) planning therapeutic tailored programs, setting realistic rehabilitation goals; iii) monitoring changes over time, also in order to evaluating the effectiveness of treatment/rehabilitation and to formulate reliable prognosis; iv) increase the number and quality of the interventions performed, with the same resources used.

Simmonds (Simmonds MJ, 1997) has affirmed "*We sometimes measure what we measure because we can measure it… we do not measure what we should measure because it is more difficult and more complex. We then use the easy measure to infer things about the difficult measure*." Of course, choosing an outcome measure is not easy and there are many factors to be considered. The first step in choosing an outcome measure is deciding what attributes are of clinical interest. There is a marked difference between the preferences of

clinical researchers/epidemiologists and those of clinicians in their selection of relevant outcome measures: the former has tended to rely primarily on self-report measures, perhaps because these are considered more global and may be easier to administer consistently within the context of clinical trials; on the other hand, clinicians on the other hand are highly reliant on impairment measures to guide treatment decisions (Law et al, 2008).

Once we determine what instruments or attributes we want to measure, we need more detailed information on validated outcome measures. Searching for literature on outcome measures is more complicated than it is for conventional clinical questions, such as diagnosis, prognosis, or treatment effectiveness. When identified an outcome measures it is essential first of all to clearly understand "what does it really measure?", for consider if it fits with our purpose; than we have to ask our self "what do we want to know and what do we want to use information for?" and finally "does this test reflect the same qualities targeted in the intervention and the qualities we wish to evaluate?". The reason for doing the assessment must guide the choice of instruments, not the availability of certain tests (Krumlinde-Sundholm, 2008).

The choice of the most appropriate outcome measures is based also on the psychometric requirements (such as reliability, validity and responsiveness) and technical and practical attributes (appropriateness, precision, interpretability, acceptability and feasibility). Knowing these properties guide in the choice, in fact, the presence of adequate levels of reliability and validity is enough for discriminative purposes (differences between subjects or groups) and predictive (classification of subjects in predefined classes for prognostic purposes), while for evaluation purposes (i.e. to detect changes over time within subjects such as in the case of efficacy analysis of therapeutic interventions) a good level of responsiveness is also needed.

Clinicians should use standardized, validated and reliable measures to evaluate their clinical decisions. Before using a measure, it should be screened to confirm that the measures have documented procedures that could be used by others and have been investigated for reliability and validity. Homegrown

outcome measures (those not standardized and/or validated) remain popular but have little value because scores for your clinic patients cannot be compared externally and cannot be interpreted by others.

However, when applying outcome measures to individual patients, clinicians need more than just information on the psychometric properties. Firstly, they need information on how to obtain the instruments, to acquire permission to use them, to administer them, and to score them and also knowledge of test administration, scoring and interpreting results for the measure to be valid and reliable (MacDermid et al, 2009).

There is increasing pressure on clinicians to evaluate their interventions for cost effectiveness in times of restrictive budgets as well as to prove their effectiveness in terms of client outcomes. Although Hanna and collaborators (Hanna et al, 2007) reported that about 50% of clinicians surveyed stated that they sometimes modified test procedures or used only selected items from standardized test procedures, awareness among clinicians about the importance of using standardized assessment tools has grown in recent years. It is also know that the systematic use of standardized clinical assessment tool is a basic requirement for collecting evidence about treatment effects. Find the right tool, which meet the needs of the clients while also being valid, reliable measures that are sensitive to change is a real challenge. In particular during development, all functions are strictly related and are influenced by each other (i.e. vision, cognitive level, motor ability) and also by many factors as age, social and cultural factors, the required task and environmental factors.

In order to choose from the different tools available, it may be helpful to use some frames of reference and basic concepts concerning test features as starting point.

1.2.2 Measurement properties

While exploring all the available tools, it is important to keep in mind that all of them could not be suitable for the assessment of an outcome. There are, in fact, two types of tools: tests and classifications.

Classification aimed to group data, subjects or objects into different categories according to common features. Classifications are useful to describe elements of a group which have similar characteristics, not necessarily to measure change.

Test means instruments that use a standardized procedure and are evaluated for their measurement accuracy of behavioral parameters within a specific population. Special test materials, score forms and detailed manual are typically provided. The test comprises several items that are typically selected through an item analysis. These items are administered in standardized way and are scored, often using a rating scale or a timed performance. The results are often reported as the sum of scores (raw scores) and successively converted to other values as percentages, age-equivalents or others. In order to be sensitive enough to detect differences between subjects or changes in the same subject, the test needs to include a range of items, from easier to those request higher abilities.

There are two main types of standardized tests: norm-referenced and criterion-referenced.

In norm-referenced tests, subjects' performance is compared to the performance of his peers and results determines how far from the mean score of the normative sample the performance is deviating. Norm-referenced tests often have a general content covering a variety of skills. These tests are intended to reflect typical performance of healthy subject, for this reason items could be too advanced for patients with developmental disabilities such as cerebral palsy; furthermore, these tests don't take in count that disabled subjects do not follow typical developmental patterns. Anyway, norm-

referenced tests are useful for identify subject's needs and eligibility for therapy, by determining the type of delay or dysfunction. Scores are commonly provided as standard deviation-based scores (standard scores, t-scores, z-scores, percentages). However, their results will not be useful for detecting changes over time (Hanna, *et al.*, 2005).

Criterion-referenced tests provide information about how a subject performs on a specific task and if he meets the fixed criteria for consider the performance successful and if it could be age-appropriated. For this aim, the performance is compared to a defined list of tasks or skills. Criterion-referenced tests measures functional skills and can be used to identify changes, since the score is based on number of passed tasks and not on the age. Typically, a rating system is used and interpretation is based on raw scores, expressed in percentage or a scaled score. If Rash analysis or Item Response Theory was used to develop the scale, raw scores are then converted to equal interval scaled scores (Rosembaum, 1998). By means of these tests, improvement in skills may be captured and the child's change over time will be measured in relation to himself (Krumlinde-Sundholm, 2008).

In addition, there are three main purposes for which standardized tests in general are constructed (Boyce *et al.*, 1991; Rosenbaum, 1998), depending on whether the measure is to be used for. There are: i) discriminative/descriptive test; used for investigating if a subject differ from others and if so, how much and for this aim, discriminative tests are often age-referenced and provide information by comparing to typical development; ii) screening/predictive measures; used to separate population between typical developing subjects and suspected or delayed ones, referring to age expectations, these tests identify subjects whom needs additional in-depth assessments and iii) evaluative measures, aimed to detect changes over time and for this reason are criterion-referenced, meaning that they are not based on the comparison with a normative sample, but on specific skills.

Although it is possible for a measure to satisfy more than one purpose, each of these functions should be validated separately.

To be useful, a test needs to fulfill some basic criteria concerning psychometric properties, as described above, furthermore, information about standard error measure (SEM) or the smallest detectable difference (SDD) is important for interpreting the change in individuals. Hanna *et al.* (2005) have created the 'Clinical Measurement Practical Guidelines for Service Providers', available on the website for the Can-child Centre for Childhood Disability Research to clearly explain the measures' psychometric properties and how to interpret them
(www.canchild.ca/en/canchildresources/resources/ClinicalMeasurement.pdf).

1.2.3 ICF framework

When choosing an assessment instrument, an organizing framework may be useful to guide the process. The purpose of conducting an assessment must be clearly expressed. The two main questions which must be clear are *what* we want to measure and *why*.
If intervention outcome is what we are interested in, the content and the target area of intervention as well as the goals for the intervention must correspond to the outcome measures chosen. If you expect transition effects on other domains of functioning, these domains need to be measured as well. When choosing an assessment instrument, different framework can be used.
The framework outlined in the International Classification of Functioning, Disability and Health framework (ICF) can be used to clarify focuses of measurements. The International Classification of Functioning, Disability and Health (ICF) is a classification system developed by the World Health Organization (WHO, 2001) which provides a framework for measuring and documenting health outcomes. The 2001 version for adults was followed in 2007 by an ICF for children and youth (WHO, 2007).
The components of ICF are (i) body functions (*i.e.*, physiologic and psychological functions of the body system) and body structures (*i.e.*, the

anatomic construct of the body); (ii) activity (*e.g.*, the execution of tasks or actions by individuals); (iii) participation (*e.g.*, involvement in life situations); (iv) environmental factors (*e.g.*, the physical, social, and attitudinal situations in which people live); (v) personal factors (*e.g.*, the particular background of individuals' life and living situation).

One interesting aspect that emerges from the ICF is the differentiation between the qualifiers of capacity and performance. Capacity refers to the subject's ability to execute a task (for hand function, e.g., to grasp, transport, or press) on the highest probable level of functioning that he/she may reach in a standardized environment (as clinical setting) that is the best possible ability upon request. Performance incorporates the real-life environment (e.g., spontaneous use of the upper limbs during activities or play), so it brings in the aspect of a person's involvement in a non-standardized environment. To simplify, capacity could mean "what subject can do" and performance "what subject does do".

Based on these aspects, we need to specify what we want to know and measure. Capacity and participation are different but equally important and it is important to pay attention to both aspects, in order to allow outcome differentiation and to guide treatment.

Clinical measures, therefore, need to be multidimensional and aimed to describe and include functioning at different levels of the ICF (Rosenbaum & Stewart, 2004).

Unfortunately, there isn't a unique tool able to include all domains of the ICF, and many authors have started to carefully analyze and code each item in the assessment according to it (Cieza et al., 2005; Gilmore et al., 2009; Krumlinde-Sundholm, 2008).

To help clinicians in the field of pediatric rehabilitation to select the most appropriate outcome measures with demonstrated reliability, validity, clinical responsiveness, and clinical utility a software called All About Outcomes© has been developed by investigators at CanChild (Law *et al.*, 1999)

1.2.4 Clinical Vs technological outcome measures

1.2.4.1 Visual assessment

The assessment of visual capabilities in the first year of life is important for monitoring the development of the infant, but also because if problems occur, they can consequently influence the whole development. One goal of child health development screening programs is to identify, as early as possible, infants at risk of future visual problems or, more generally, developmental disabilities, in order to make an early intervention possible and incisive (Morante et al, 1982). Visual assessments of infants are not easy: many assessment techniques that are used with adults are based on the patient's response thus not suitable for infants since they have such a short attention span. Nevertheless, in recent decades it has become possible to use new assessment methods that have been adapted for the needs of younger patients, that require neither the collaboration of the subject nor special abilities of the examiner, so they are suitable for very young infants, non-verbal or non-collaborative subjects (e.g. patient with mental delay or behavioural problems). Abnormalities of visual function, which are common in infants with or at high risk for developmental disabilities (in particular cerebral palsy), may include poor visual acuity, reduction in visual fields, disorders of eye movement, strabismus, and complex visual-perceptual defects.

These abnormalities are mainly identified through behavioural assessments and their description is often related to the ability of the operator in understanding and interpreting the subject's response.

The more used clinical assessments are presented (Chorna et al, 2017):

Optokinetic nystagmus, an involuntary motor reflex responsible for the ability to follow a moving object while keeping a steady gaze, is tested using a rotating a striped drum, black and white patterned board, or mirror. The distance between the infant and the drum at cessation of the reflex response is recorded.

Visual field assessment is performed using confrontation techniques, during which an examiner attracts the infant's attention centrally. An assistant introduces another object into the periphery, and the examiner determines if the infant orients to the peripherally presented object, estimating the visual arc. A variant of this, kinetic perimetry, utilizes two white spheres as the objects.

Oculomotor function assessments evaluate the presence of expected or atypical eye movements such as light gazing, light fixation, gaze fixation, and tracking.

Visual acuity of cooperative infants without severe central visual impairments can be assessed using preferential looking toward patterned cards (Teller, Keeler, Cardiff, Griffiths, Berkley, and variations). Acuity is measured in cycles per degree, and recorded as the threshold of the smallest width of the grating on presented cards that elicit a typical response as referenced to normative data. Visual acuity can also be tested using optotypes in children as young as three years, but none are designed for infants under two years.

Based on this, it emerged how the clinical assessment is strongly related to the evaluative context in terms of stimuli presentation, response's recording, required time, infant's attention and in consequences to the operator whom conducted the test. It is important also to highlight that infants with developmental disabilities could present, together with the visual impairment, also postural problems and gaze movements should be assessed also in relation to head movements. In addition to that (with exception for the visual acuity) the clinical assessment is often made with objects with similar features, but not standardized, as should be done for an assessment tool. Finally, to my knowledge, there are not many clinical centres in which there are trained therapists able to conduct a visual assessment. The possibility to optimize and standardize the assessment of gaze movements, connecting eye and head movement, could represent a new panorama in this field.

1.2.4.2 Assessment of upper limb movement

Clinicians interested in evaluation and treatment of the upper extremity can choose from a wide range of assessment tools and classification systems. Both types of instruments may be used for outcome assessment in multiple domains. Choosing the correct instrument may be a complex undertaking and should be based on a variety of factors, because even well-developed and validated measures may not of help in tasks for which they were not specifically created and validated. This process is greatly facilitated by knowledge of instrument content, methodology, and clinical use.

Here a brief description of the available clinical tools, most used for children:

- The *Manual Ability Classification System (MACS)* is a questionnaire which describes how children with cerebral palsy use their hands to handle objects in daily activities. MACS describes five levels based on the children's self-initiated ability to handle objects and their need for assistance or adaption to perform manual activities in everyday life. It is suitable for children between 4 and 18 years and the interpretation of the level is related to the age. The MACS classification has reported good validity and an excellent inter-observer reliability among health professionals, with an intra-class correlation coefficient (ICC) of 0.97 (http://www.macs.nu).
- The *House Functional Classification System (HFCS)* is a nine-level functional classification system which provides categories to describe the role of the assessed hand in children with cerebral palsy as a passive or active assist in bimanual activities. It was designed to provide an evaluation of the function in the affected hand for patients with thumb-in-palm deformity due to spastic hemiplegic cerebral palsy, to evaluate children pre and post surgery and to establish the efficacy of botulinum toxin injections. In 2008, Koman et al. (Koman et al, 2008) introduced the *Modified House Functional Classification (MHC)* with the intention of improving the score discrimination of the original classification, to make it better suited for monitoring patients and to evaluate

treatment efficacy. MHC involves the observation of patients during the performance of increasingly complex functional activities

In the Modified version, the MHC concerns the performance of increasingly complex functional activities and evaluates three object-related basic categories of hand action: reaching, grasping and manipulating objects, which are used in combination to perform activities of daily living.

- The *Bruininks-Oseretsky Test of Motor Proficiency* is a standardized, norm-referenced measure used by physical therapists and occupational therapists in clinic and school practice settings. It was revised as the Bruininks-Oseretsky Test of Motor Proficiency, Second Edition (BOT-2; Bruininks & Bruininks, 2005). The *BOT-2* is a widely adopted motor proficiency test used to measure fine and gross motor skills of children and youth, between the ages of 4 and 21. It is used to evaluate motor performance, i.e. fine manual control and manual coordination, and it has a subtest and composite structure that highlights motor performance in the broad functional areas of stability, mobility, strength, coordination, and object manipulation. It investigates the ICF activity domain. It has satisfactory reliability coefficients: test-retest reliability scores averaged 0.87 (unspecified statistic) for the complete battery. Inter-rater reliability varied between 0.90 and 0.98 (unspecified statistics) (Deitz et al, 2007).
- The *Jebsen-Taylor Hand Function Test (JTHFT)* is a short test used to assess a broad range of uni-manual hand functions required for activities of daily living. It evaluates adults with neurologic or musculoskeletal conditions involving hand disabilities. Each item is performed with each hand separately, the non-dominant hand being used first. The test assesses the speed, not the quality, of the performance and it is approximately 6 to 30 min long. However, it can take up to 45 minutes with slower patients. It can be carried out in children ranging from 6 to 18 years. Psychometric data for children with hemiplegia however are not available. According to ICF the JTHFT investigates the activity domain.

- The *Box and Block Test (BBT)* is a rehabilitation measure used to assess unilateral gross manual dexterity. It has been validated for children and adults with acute and chronic stroke. It can be use with children ranging from 6 to 12 years and it interest the ICF activity domain. The test is 5 min long and it is carried out thanks to the use of special equipment that is a rectangular box that divided into two squares of equal dimension by means of a partition and 150, 2.5 cm wooden cubes. The individual is instructed to move as many blocks as possible, one at the time, from one side of the partition to the other for a time of 60 seconds. It has an excellent test-retest reliability (ICC =0.95) and interrater and intrarater reliability (r=0.95) in spastic hemiplegia.
- The *Assisting Hand Assessment (AHA)* is a standardized, criterion-referenced and performance-based test for use with children aged 18 months to 12 years, with hemiplegic cerebral palsy. Additional versions are available: *Mini-AHA* for infants aged from eight to 18 months old and recently also the research version *Assisting Hand Assessment in adolescents (Ad-AHA*) for adolescents and adults. It evaluates the spontaneous use of the assisting hand (affected hand) during a semi structured play session with specific toys (from the AHA test kit) requiring bimanual handling. The intent of AHA is to reflect how the person usually acts (performance), not the best possible capacity. The test investigates the ICF activity domain. The AHA is video-taped in a standardized approach, and the play session lasts about 20 minutes. A certified rater scores the video recording of each child's performance on 20 items, grouped in 6 main functions of the two upper extremities, according to a four-point scale. Inter-rater, intra-rater, and test-retest reliability of the AHA have high ICC (0.99 and 0.98, for the small- children and for the school- children's version, respectively (Holmefur et al, 2007).
- *ABILHAND-Kids* is a measure of manual ability for children with upper limb impairments**.** The scale measures a person's ability to manage daily activities that require the use of the upper limbs, whatever the strategies involved. It has been validated in children with cerebral palsy from 6 to 15 yrs and was calibrated according to the difficulty children encountered when performing

manual activities as perceived by their parents. The short condition-specific questionnaire measures 21 mainly bimanual daily activities and refers to the activity domain of the ICF. The ABILHAND-Kids has excellent clinical use. It is available free online and is complete with an online analysis of the data. It is quick to administer and is completed by parents who offer a different perspective on children's manual ability in their everyday life. Furthermore it has a high reliability (R = .94) and a good reproducibility over time (R=.91) (Gilmore et al, 2010).

- The *Children's Hand-use Experience Questionnaire (CHEQ)* is a measure designed by occupational therapists to evaluate how children with unilateral hand dysfunction use their affected hand in everyday bimanual activities. The CHEQ has recently been adapted to an online version with a four-category rating scale and consequently new psychometric testing was warranted. The clinical utility of the updated CHEQ is strengthened with the free online version, which has a practical administration time of approximately 30 minutes and allows efficient generation of reports. The CHEQ is a useful tool as the items are frequently carried out by children in their daily routines, are not sex- or culture-specific, and are not strongly dependent on other functions such as gross motor or cognitive function. The CHEQ differs from other upper limb assessments in that the scales capture not only a child's perceived grasp efficacy and time taken to complete the activity, but also considers the experience of feeling bothered by impaired hand function. The validation sample (n=242) was heterogeneous, with participants from six countries worldwide. The excellent test–retest results, with intraclass correlation (ICC) values ranging from 0.88 to 0.91, are a major strength of this tool. It interests the activity domain of ICF.
- The *Modified Ashworth Scale (MAS)* is a scale which measures the impairment of body function according to ICF. It is based on the scale proposed by Ashworth (1964) but Bohannon and Smith added an extra item at the lower end (grade 1+) in 1987. It is used in children between 6 and 12 years. The intra-

rater reliability of MAS is from 0.55 to 0.83 and the inter-rater reliability from 0.45 to 0.84 (Morris et al, 2002). However, in recent studies in children with CP the inter-rater reliability was described as low (Mutlu, 2008). In the latter the authors concluded that an assessment of spasticity using MAS was not very reliable and that it should be used with caution.

- The *Quality of Upper Extremity Skills Test (QUEST)* is a descriptive, impairment-based and evaluative measure designed to evaluate the quality of upper extremity function, separately for each arm, in children between 18 months and 8 yrs of age with spasticity. The assessment is conducted in a play environment. However, the QUEST is reportedly not playful and engaging, especially for children who are younger and have more significant impairments. The test is from 30 to 60 min long. According to the ICF, the QUEST measures the child's capacity in the domain of activity with the grasps subscale, and the impairment of body function with the dissociated movement subscales. Variable intra-rater reliability (ranges from 0.51 to 0.96 with all coefficients except one greater than 0.70) and strong test-retest reliability (ranges from 0.75 to 0.95) have been reported (Klingels et al, 2010).
- The *Melbourne Assessment of Unilateral Upper Limb* Function (MUUL) is a quantitative test of quality of movement in children with neurological impairment. It values the children aged from 5 to 15 years with neurologic impairments. In 2008 a modified MUUL for children from 2 to 4 years of age was developed (*Melbourne Assessment 2 MA-2*). MUUL is a criterion referenced test based on 16 items scored on a 3- to 5-point ordinal scale and comprising tasks that are representative of the most important components of unilateral upper limb function (reaching, grasping, releasing, and manipulation). MUUL measures both capacity in the activity domain and some aspects at the body functional level e.g. range of motion, fluency). The child is recorded on a video and subsequently scored. MUUL investigates the ICF body function domain. These findings indicate that the Melbourne Assessment of Unilateral Upper Limb Function is a reliable tool for measuring the quality of unilateral upper-limb movement in children with CP. However, some studies

have suggested that the MUUL might not be sufficiently sensitive. Results revealed high inter rater reliability (0.95) and intra rater reliability (0.97) for total test scores. Test–retest results revealed moderate to high intra rater reliability for item totals (mean of 0.83 and 0.79) for each rater and high reliability for test totals (0.98 and 0.97). These findings indicate that the Melbourne Assessment of Unilateral Upper Limb Function is a reliable tool for measuring the quality of unilateral upper-limb movement in children with CP. However, some studies have suggested that the MUUL might not be sufficiently sensitive (Randall et al, 2001).

As seen here, there are several available clinical tools for the assessment of upper limb function, even it is possible to identify some limitation and aspects which could be improved. Firstly, when using a questionnaire, answers could not perfectly reflect subject's abilities, in particular when parents fill in the answers, because they are influenced by many factors, like the perception of their children, their emotional mood and feelings, the moment in which they answered to questions (often in waiting rooms while their children are doing therapy). Another consideration is that the functional global level of upper limbs could not emerge in a clinical assessment and clinical score sometimes doesn't reflect the real use and ability of the upper limb: it could happen that children with unilateral cerebral palsy are aware of the goal of their evaluation session, so they give their best to complete the test, but this is not spontaneous and in their life they demonstrate a lower use of upper limb; vice versa subject could be inhibit during the assessment resulting less able than what they usually do. On the other hand, there are assessment tools. These are often conducted in a playful environment, and children were asked to perform some actions. Those assessments need to be administered in most cases with the specific material of the test and there are often strict instructions about the available time. Moreover, there are few tests which require a certificated training course for operators (e.g. the Assisting Hand Assessment), but in many cases the administration of the test is free from a qualifying examination (e.g. Melbourne Assessment 2, Box and Block test..); this point highlights the

important correlation between the experience of the therapist and the results of the test session, together with the consequent scoring. Another point is the movement analysis: many tests allows to describe upper limb movement from different perspectives, but they give information about the range of movement in terms of possible positions of articulation, in some cases they are focused on the success of a goal directed action (sometimes independently from the execution) and the global analysis of the movement in terms of quantity of activity of each upper limb, cooperation between arms and the quantitative relation in the use of the two upper limb couldn't result very accurate.

1.2.4.3 General considerations

Based on all mentioned factors, choosing the most appropriate outcome measure it is a crucial process both for clinical and research practice. It is important to take into consideration the several clinical measurements available so that the best one may be adopted for the assessment needed. There are several clinical outcome measures, many of them are also standardized and already validated on a population and this process is always ongoing and upgrading. In some cases, there is also a training course which enable to administer the test, but in many cases, any therapist could use it and this raises the strong correlation between the right application, the score process and the experience of the operators. This is related to inter- and intra-rater reliability, which is one of the requirements of a good outcome measures. Another important property of a test is the repeatability in terms of test–retest reliability, that means that there is no (or minimum) variation in measurements taken by a single person or instrument on the same item, under the same conditions, in a short period of time. In addition, highly related to test-retest, there is the smallest detectable change, that is the minimum change of the score that the outcome measure could detect and can be interpreted as a real variation; it

guides the choice of the outcome measure and the study protocol to design and follow.

Available measures are mainly clinical, but in the last years, the progress in technology and engineering has interested also the assessment of different aspects of the human development. Thanks to all the knowledge about the properties that a "good outcome measure" must have and mainly about its limits, new technological tools have been developed, with the purpose of reducing the lack of quantitative and reliable measures in clinical practice. In addition to that, an automatized and standardized way to measure an aspect could reduce all issues related to the operator whom conduct and/or score the test, increasing inter- and intra-rater reliability and overcome the limitation of applicability of tests, related to the smallest detectable changes. Within this frame, two main lack of quantitative and highly reliable outcome measures of young subject have been highlighted: i) the visual assessment of infants and ii) upper limbs movements assessment of children in school age.

Thanks to the contribution of technology, there are some innovative devices which can be used for quantitatively measure some parameters and integrate the clinical assessment conducted with traditional tools. Some researches started to use commercial devices, trying to understand if they could be suitable also for a clinical purpose. One example could be represented by the use of eye trackers, born for the recording the drivers' attention and avoid sudden-onset sleep (in particular, the SmartEye system, http://smarteye.se/) or the growing industry of wearable sensors, that is, for example, the accelerometer which can be wear as bracelets (some products are: Axivity, FitBit, Actigraphs). The use of the accelerometer has increased considerably in recent years because, alongside traditional scientific and aerospace applications, it has been adopted in many civil fields (smartphone, test, mechanics, gambling..) often embedded with other sensors as gyroscopes, magnetometers and so on. With this increase of applications, many tools have been diversified and today there are many different available types, each with different functional and constructive features, many of them also with a reasonable cost.

However, choosing and using a technological tool is not immediate because first of all, it is crucial to select the most appropriate one, basing on the clinical/research needs and properties of the tool (application methodology, required knowledge…).

The use of technological devices is increasing in scientific literature; but if on one hand these tools are very useful thanks to the fact that they record and provide raw data, which can be elaborated with dedicated analysis in order to get the required values; on the other hand, they need to be validated before being used for the purpose of the study. In the validation process, the selected tool must demonstrate its suitability and applicability for the purpose, providing normative data. This process is indispensable and must proceed the study and in some cases, it could not provide the waited result.

Technological outcome measures can provide quantitative data, but another aspect to be considered in collecting and analyzing those data is that they are often designed for "standardized applications", which means that the quantity and quality of data are related to the methodology of use.

In evaluative practice, the assessment can be done in three situations: structured, semi-structured and completely free. Each type of situation provides different data: the highest reliability is obtained in a structured situation, because additional casual or wrong data are avoided by a repetitive trial, but it could not be easily applicated with infants and when it is possible, it could not reflect the real performance because the subject's behavior is completely guided by the operator. On the other hand, a free situation, the most ecological setting for a child, in which he is free to behave as usual, make the acquisition of the data and their elaboration very complicated, in relation to the variety and quality of collected information because together with the target data, there are a lot of other data which has to be removed in the data processing.

Semi-structured situation often represents a good compromise, but in terms of new technological tools the ideal situation must be defined and sometimes several tentative are required.

Subject's age is one of the first constraint while choosing the situation of the assessment, because a young infant (in the firsts months of life) does not follow the indication about how to behave and the moment of the recording is not often the moment in which he shows the required movement.

This make the application to infant and children challenging because their behavior is not easily predictable, and, for example, the required movement is done while doing other movements, or in a too long time, or being distracted by something or together with other behavior. The high reliability of a technological measure is related to a reliable application and this is the real challenge when combining clinical and instrumental evaluation with infants and children.

Of course, the real challenge and the best opportunity is represented by the possibility to record the highest quantity and quality of data source of interest while subject is free to behave in the most natural and spontaneous way.

1.3 Purposes of the thesis

In this scenario, a growing number of new technological techniques have been developed, to get insights into the understanding of infants' and children's development. This knowledge is, and it will be, useful to standardize the clinical assessment and increase the reliability of data and also to potentially equate the assessment in all clinical centers, without be totally dependent to the therapist, whom could be more or less expert than others. Another possible application is represented by the tele-rehabilitation, that is the possibility to remotely follow and customize a rehabilitative intervention from the clinical center. Tele-rehabilitation frontier permits to clinicians to provide intervention to many subjects simultaneously, and to patients' families to avoid the transfer to clinical center allowing a home-based intervention which could be done in the best moment of the day, making it easily effective and intensive, and

centered to the family whom actively participate to the training remotely followed by clinical staff.

Starting from these findings, in my PhD period, I applied technology for two main purposes: i) to improve infants' and children's assessment and ii) to propose innovative rehabilitative proposal, mainly based on technological tools.

As a therapist, I am very interested in finding new assessment techniques that allow to make the assessment more qualitative and reliable and complete the traditional clinical practice, adding evidence on treatments. In addition, the use of technological devices in rehabilitation could improve its quantity and quality, by making it more quantitative, donating the possibility to be done at home, involving families, increasing the customization of activities and permitting to the therapist to follow many patients simultaneously (thanks, e.g. to tele-rehabilitation briefly descripted above).

The encounter between clinical purposes and technological solutions has been possible thanks to the collaboration between IRCCS Fondazione Stella Maris and the BioRobotic Institute of Scuola Superiore Sant'Anna and in particular to the joined laboratory born from their partnership called Neurodevelopmental Engineering, in which I developed and conducted this work of thesis.

Data reported in this dissertation are largely based on published papers that are the scientific products of the PhD program.

Chapter 1 introduces the neurodevelopmental disorders and the importance of their early detection and in relation to this, the importance of outcome measures. After an overview on the clinical available outcome measures, the main characteristics of a tool and a brief exploration of visual assessment and assessment of upper limb movements are presented.

Chapter 2 presents the design, development and preliminary testing of a new system for measuring infant's gaze in the wide visual field called CareToy C. The idea is to understand if an Eye tracker quite never used before for the assessment of infants, could be adapted to this purpose and how, integrating the traditional visual assessment.

Chapter 3 is focused on the application of CareToy H platform, that is a biomechatronic gym aimed to promote infants' development: it has been used for measuring manipulation capabilities and for rehabilitation of infants at risk for Cerebral Palsy. Its application is hypothesized to move also to other population and for this reason a small sample of infants with Down Syndrome has also been tested with CareToy.

Chapter 4 presents the technology applied to the measurement of upper limb use of children. A literature review introduced the use of sensors for the assessment of upper limb and a validation of Actigraphs was conducted.

In **Chapter 5**, an innovative approach is presented, Action Observation Therapy. Its use is described with a systematic review and then its combination with Information and Communications Technologies is showed in a RCT called Tele-UPCAT, from which a pilot study is extracted and presented.

Summary and conclusions are presented in **Chapter 6**.

CHAPTER 2

CARETOY C: A NEW SYSTEM FOR QUANTITATIVE EVALUATION OF INFANT GAZE CAPABILITIES IN A WIDE VISUAL FIELD

2.1 Introduction

The visual assessment is crucial in children, in particular in the first year of life, when is important for monitoring the development, but also because if problems occur, they can consequently influence the whole development (Dubowitz et al, 2005).

One goal of child health development screening programs is in fact to identify, as early as possible, infants at risk of future visual problems or, more generally, developmental disabilities, in order to make an early intervention possible and incisive (Morante et al, 1982).

However the visual assessment of infants poses specific challenges: most of the techniques that are used on adults are based on the patient's response, and are not suitable for infants.

In recent decades it has become possible to use new assessment methods that have been adapted for the needs of younger patients, that require neither the collaboration of the subject nor special abilities of the examiner; so they are suitable for very young infants, non-verbal or non-collaborative subjects (e.g. patient with mental delay or behavioral problems).

In recent years, eye-tracking has become an increasingly popular tool among researchers in the neurodevelopmental field (Gredebäck et al, 2010). Significant advances, particularly in the area of automated eye-tracking, have made this technology much more user-friendly and applicable to human infant populations than ever before (Corbetta et al, 2012)

Despite this fast growing interest in infant eye-tracking, the use of this technology has been mostly applied to capturing infant eye movements and gaze patterns when looking at objects or scenes depicted in two dimensions on a computer screen (Quinn et al, 2009; Johnoson et al, 2003; Farzin et al, 2010; Farzin et al, 2011; Gredebäck et al, 2004: Johnson et al, 2004).

Significant advances in the eye-tracking have made this assessment of infant visual capabilities easier, however, eye-tracking still requires the subject's collaboration, in most cases and thus limiting the application in infant research. Focusing on the assessment of visual attention, the disengagement of attention is the most studied function (engagement, disengagement and attention shift) by using gap and/or overlap paradigms. In both cases there are two visual stimuli, one on the midline and the other one the on periphery. Both paradigms study the infant's gaze shift from the central to the peripheral stimulus. In the gap condition, the midline stimulus disappears before the peripheral one appears, while it remains present in the overlap paradigm (Blaga and Colombo, 2006; Atkinson et al, 1992) . Clinical studies, without using an eye-tracker, have shown that infants in their first year of life need more time to shift the attention from a central to a peripheral stimulus in the overlap condition (Guzzetta et al, 2001). Moreover, this effect is more evident in younger infants and decreases with age; this allows us to hypothesize a maturational effect on the gaze shifting competence.

Another interesting aspect is the association between visual and auditory stimuli.

When they are combined, the response latencies are fastest with audio-visual targets than the visual targets alone (slower), or the auditory targets (slowest) (Neil et al, 2008). All the paradigms have been applied in a limited visual field. Moreover, there is a lack of transferability to clinical practice, and thus it emerges the need for a new tool to measure the paradigms and explore the most common visual competences in a wide visual field. In this section, the system called CareToy C is presented (CareToy for Clinics). CareToy C and the specific software for the management of the audiovisual stimuli would allow the clinicians to develop clinical tasks for the quantitative assessment of visual attention i.e. time and degree of gaze (head and eyes components, separately) in all three (gap, overlap and multisensory) paradigms in a wide field condition.

2.2 Materials and methods

On the basis of the clinical requirements, the following features have to be taken into account. The system should be able to:

- measure wide range of visual field (120°);
- provide audio-visual stimulation;
- identify five different areas of the visual field for audio-visual stimulation covering 120°
- measure infants gaze i.e. single contribution of eyes and head respectively;
- represent a non intrusive setup and an ecological approach during clinical trials;
- be adapted according to the infants' needs (aged 3–12 months).

Hardware

The system consists of an eye-tracker integrated into a customized mechanical structure. The five points are placed in specific positions: one in the center, two on the right side (30° and 60°) and two on the left side (30° and 60°). In order to provide audio-visual stimulations, each point of interest comprises a screen (10.5″) and a speaker. A mechanical structure has been designed in order to fix the screens and the speakers at specific distances and angles (Fig. 1). The screens active area represents our area of interest (AOI), so our stimuli dimensions are 220 × 129 mm. We selected USB-VGA converters and an external audio device for managing multimonitors images and sounds signals. Custom software has been developed for the management of the audio-visual stimulations. A dedicated PC has been assigned for the management of these audio-visual stimuli. The external sound card device with six outputs is used to send the sound to the five speakers; the sound card communicates with the PC via the USB port and is connected to the speakers with RCA cables. The desktop is extended to five screens using the USB/VGA converters (Fig. 1). As far as the eye trackers are concerned, the technological background of the

eye tracking technology has been studied to choose the best solutions for our application. In order to obtain a non-intrusive system for measuring the infant gaze, video-oculography (VOG) has been selected. This family of eye trackers is based on corneal-reflection. They assess a video input of the pupil's highlight reflected on the cornea, usually from a light source invisible to humans, in the infrared range of the spectrum. The center of the pupil and the corneal reflection are tracked in real time, providing information about the participant's point of gaze (POG) on the stimulus. Among the available eye trackers we selected the SmartEye system with six cameras running at 60 Hz due to the following main features:

Most flexible cameras placement in our setup One of the main advantages of the SmartEye system is its flexibility. The cameras can be positioned independently one from the other, but each camera should be oriented in such way so that the subject's head is the camera's point of focus. In addition, the number of cameras is not fixed but can be adapted in order to cover the required visual field (±60°).

Gaze and head tracking The system uses IR-diodes to illuminate the face of the subject in order to minimize the effect of varying lighting conditions and use the reflections of these IR flashes on the cornea ("glints") to find the center of the eyes, rather than other systems where the eye center is estimated using the head model (Fig. 2). This allows a more accurate identification of gaze direction with fewer errors in the head pose estimation. This feature, i.e. the possibility to measure both eye and head components is extremely important since the visual field proposed in this work is wide and each movement of the infant to reach the audio-visual stimulus is composed of both components. Finally, in order to complete the system and to adjust the setup according to the infant's needs, a seat has been purposely designed allowing for adjustments to the height (the infant's eyes should be at the height of the screen's center) and the distance from the screen (≈60 cm). In addition, two different accessories complete the seat that allow us to position infants that have not yet

reached the stage of torso control, as well as infants that can maintain the sitting posture independently (fig 3).

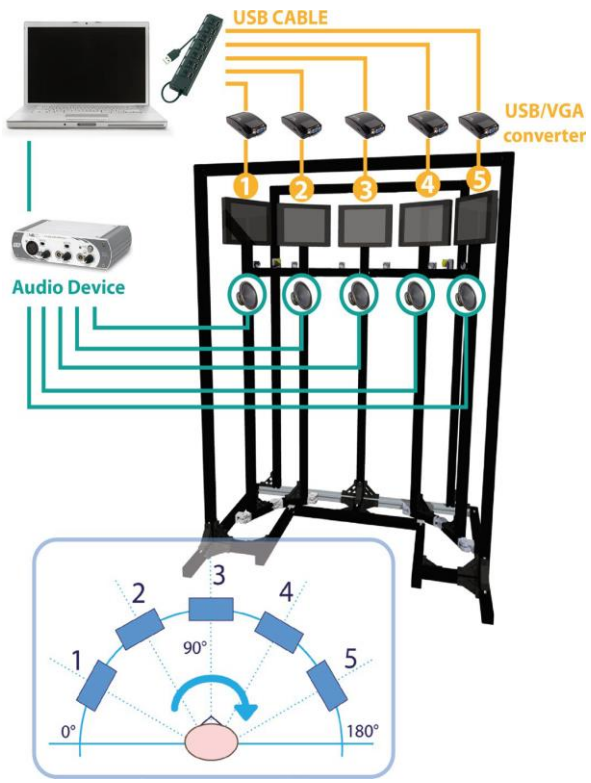


Fig. 1 Schematic system overview. The mechanical custom structure represents the support for the five screens, the five speakers, the six SmartEye cameras running at 60 Hz and two IR-diodes used for illuminating the face of the subject in order to minimize the effect of varying environmental lighting conditions and for using the reflections of these IR flashes on the cornea ("glints") to find the centre of the eyes. The stimuli management has been obtained using a laptop combined with the audio–video external devices. In the *lower part* of this overview, it is possible to observe the gaze heading frame of reference

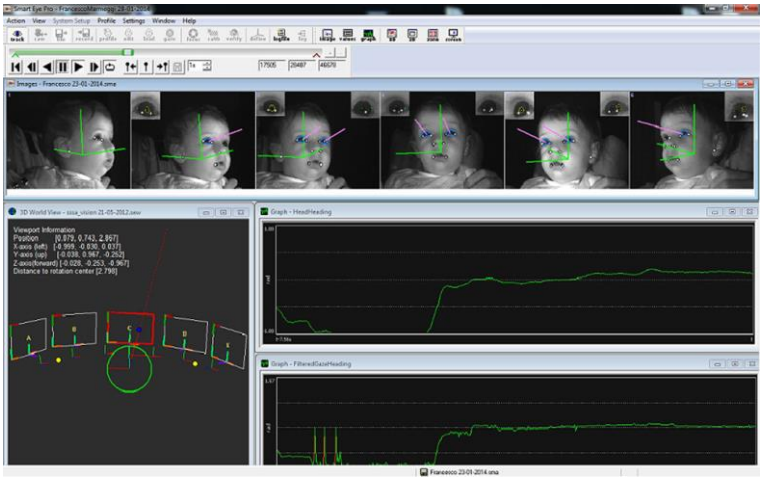


Fig. 2 SmartEye Graphical User Interface (GUI). In the *upper bar* it is possible see the *pink vectors* that represent the infant's gaze vector; in the *lower* window on the *left*, there is the 3D representation of the external 3D setup with the visualisation of gaze intersection on a object modelled in the 3D world (i.e. the gaze vector intersects screen n.3) and on the *right* there are the typical gaze heading and head heading signal profiles

Fig. 3 The final version of the system

Software

Purposive software has been developed to create sequences of stimuli (images and sounds) and send them to the screens and speakers. Through the GUI it is possible to set some parameters: (a) the total number of images composing the audio-visual tasks sequences; (b) the stimuli (images and sounds) to be presented on each screen and (c) the duration of each stimulus (each sound and image can have a different duration). SmartEye System computer desktop and the PC that manages the presentation of the stimuli are connected to a Local Area Network (LAN) and they synchronize their local time using the Network Time Protocol (NTP).

Experimental paradigm

The infant is placed at 60 cm from the screens and the environment is black to avoid distracting factors and to allow the infant to devote their attention to the audio-visual stimuli.

The stimuli were chosen in order to be interesting and attractive for infants in their first year of life. There are two series of images for two main categories of age: geometric, circular, concentric or high contrasted pictures (like black/white chessboards, concentric black, white and red images) for younger infants and human faces for older infants. We designed three different tasks to test our hypotheses:

- "attention task" (AT) is the reproduction of the gap paradigm: the stimulus is presented on the central screen; after 3 s this stimulus disappears and a different stimulus is proposed on one of the peripheral screens (30° or 60° on the left/right side, randomly); we expected that the system would be

able to measure the shifts of gaze in a non-competitive situation both at 30° and 60° on the right and on the left side of visual field.

- "fixation task" (FT), is the reproduction of the overlap paradigm: the stimulus is presented on the central screen; after 3 s a lateral stimulus simultaneously appears on the peripheral screen (30° or 60° on the left/right side, randomly) in competition with the central one; we expected that the system would be able to measure the shifts of gaze in a competitive situation both at 30° and 60° on the right and on the left side of visual field.
- "audio-visual task" (AVT), the stimulus is presented on the central screen; after 3 s an audio stimulus is presented simultaneously with a visual one on the peripheral screen (i.e. spatially and temporary coherent); we expected that the system would be able to simultaneously produce the audio and visual stimuli and that it would be able to measure the same parameters of AT, detecting the differences in the time parameters between the two tasks (AT and AVT).

The basic sequences of each task were repeated in order to obtain a global duration of 100–120 s. These durations were calibrated based on the average time of visual attention in the first year of life.

Sample

Nine healthy, born at term, infants (4 males, 5 females) with an age range between 4 and 10 months (mean 7 ± 1.73 months) were assessed using this system at the IRCCS Fondazione Stella Maris. Each infant performed three different tasks described below (attention, fixation and audio-visual) in a random order. This clinical trial has been approved by the Ethics Committee of Pisa University Hospital and the Tuscan Region Pediatric Ethics Committee (Italy).

Parameters of interest

Ideally, the profile of the Gaze signal is represented by a ramp divided into three phases (Fig. 4): (a) a fixation phase on the central screen, (b) a gaze

movement in the direction of the peripheral screen and (c) a new fixation on the peripheral screen. From this profile, it is possible to identify three main time instants:

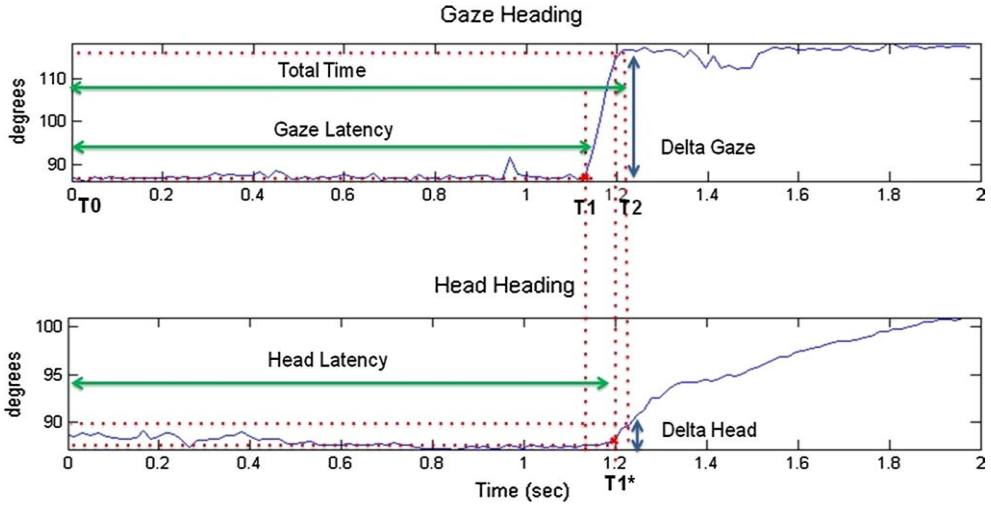


Fig. 4 Gaze heading and head heading signal profiles. An example of gaze heading and head heading time response with relative Total Time (TT), Gaze Latency (GL) and Head Latency (HL) parameters

- T0 the stimulus appears on the peripheral screen
- T1 the gaze begins to move towards the peripheral screen
- T1* the head begins to move towards the peripheral screen
- T2 the gaze intersects the peripheral screen and the movement is completed

Starting from the previous time instants the following time parameters have been selected:

- T0–T1: Gaze latency (GL)
- T0–T1*: Head Latency (HL)
- T1–T2: Move time (MT)
- T0–T2: Total time (TT)

In addition, we identified delta parameters related to the angular movements:

- Delta Gaze (DG) the angular shift of gaze during MT
- Delta Head (DH) the angular shift of head during T1*–T2 interval

- Delta Eye (DE) the difference between DG and DH

Starting from a more accurate literature evaluation, it is important to take into account several factors influencing data quality: different participants' characteristics, level of the operator who performs data acquisition, kind of task used during trial sessions, varying environmental conditions, and last but not least, eye tracker specifications, in terms of: cameras resolution, sharpness of the eye images and calibration procedure. We carefully took care of this last aspect in order to prevent a high data loss rate. Thus, at the beginning of each trial session, we carefully performed a double calibration in order to calibrate both the system and the infant's gaze. The system calibration was divided into two phases: the first one aimed to adjust the six system cameras to desired positions (small displacements were needed in order to keep the infant's face centred in at least four cameras) and relative camera brightness
and focus. This phase consisted of showing attractive and coloured figures combined with different sound stimuli to the infants, sequentially in the five screens. Then, in the second phase, SmartEye Pro 5.9 application automatically detected the current positions and orientation of the cameras, by moving a little chessboard in front of the cameras to calculate their relative position in respect to the whole setup. Starting from these positions, it calculates head and eyes positions. The purpose of the gaze calibration was to determine the difference between the visual and the optical axis of the eye. We defined five calibration points corresponding to each centre point of the five screens as objects in the syntax of the 3D world model, and the system automatically created calibration points on each screen. This procedure consisted of showing a smile emoticon growing from the centre of each screen combined with a sound stimulus in order to address the infant's gaze to the AOI centres. If the calibration was not successful, e.g. one or more calibration point was missing, we repeated the process at least two times to obtain a satisfactory calibration for all five locations. Thanks to SmartEye Pro 5.9 application, it was also possible to quantify the accuracy of the calibration points obtained during recording (Fig. 5). For each point we checked the accuracy and standard deviation of the

calibration. The accuracy depended on various parameters such as: (a) distance from the cameras, that we assumed remained fixed (60 cm); (b) distance and position of the screens (fixed at 60 cm); (c) individual differences among infants that we did not include in this study. The calibration was repeated until all calibration points had accuracy lower than 1°. Moreover, regarding data quality and according to SmartEye technical specifications, it was possible to perform a preliminary data quality evaluation thanks to an important system parameter: *Gaze direction quality* that is presented in the range (0.0, 1.0). When the *Gaze direction quality* is 0, it means that the system is not tracked correctly and loses the infant's gaze. This parameter expresses a threshold value in order to distinguish reliable data from that to be rejected. Basically it represents a trade-off between recorded data system availability and data accuracy. In order to maintain a reasonable amount of data with sufficient accuracy, we decided that values of gaze direction quality lower than 0.4 would be considered unreliable and consequently rejected. Furthermore, analyzing data in comparison to the gaze heading frame of reference
(Fig. 1), we decided to exclude the following cases from the analysis:

- all the trials in which the infant's gaze did not start from the central stimulus
- all the trials in which the gaze shift was outside the working space (±90°).

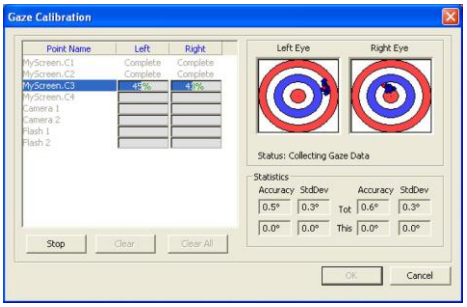


Fig. 5 Gaze calibration procedure. The *red dot* shows where the current un-calibrated gaze intersects a plane, orthogonal to the current world point and a vector pointing towards the centre of the eye. The *blue dots* represent all saved samples, whereas the *green dots* show the samples once the calibration algorithm has been run on them. Ideally the *green dots* should be in the middle of both the target. One *circle* in the target corresponds to ±2° of accuracy. We carefully checked that the *blue dots* were close together without outliers, and any outliers that were found were

cleared and new samples were added again. We manually repeated this operation until the noise became smaller.

Data analysis

Data analysis was devoted to the identification of the time parameters mentioned above and the relative angular movements of head and eyes. T0 can be easily extracted by the time stamp corresponding to the instant in which the stimulus appears on the peripheral screen given by the stimuli management software. The acceleration vector corresponding to *Gaze heading* has been used to find T1. When the gaze begins to move, there is a peak in the acceleration signal and this represents the exact parameter of interest (Fig. 6). The acceleration vector has been obtained starting from *Gaze* vector (second derivative of *Gaze* parameter).

The adopted strategy was to find all the highest points that represent the peaks and relative time coordinates, and rejecting all gaze movements in the wrong direction. The value of the *Gaze heading* at the left of the subject is 0°, 90° at the centre and 180° at the right (see the Fig. 1). If the infant shifted the gaze from screen 3 towards screen 4, we expected a growing gaze angle (for opposite movements, we expected decreasing values). The same procedure was used to find T1* from the head heading vector. T2 can be selected from the SmartEye software. Starting from a 3D representation of the external world, the software returns information about the intersection of the gaze with a screen, so it is possible to exactly determine the time in which the gaze intersects each peripheral screen (T2). The spatial parameters of the angular shift are calculated accordingly.

Clinical data was analysed by means of the Statistical Package for Social Sciences (SPSS, version 20.0). Means and standard deviation were calculated and reported in Table 1.

Mann–Whitney U independent sample test was used to analyse the following comparisons

in all three tasks: differences between 30° and 60° for each parameter and differences between DE and DH both at 30° and 60°.

The same non-parametric test was used to compare differences between attention and audio-visual tasks (both 30° and 60°) for each time parameter.

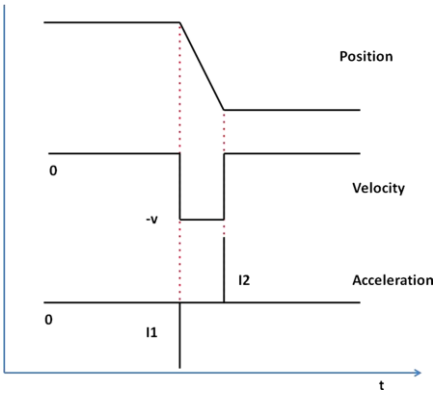


Fig. 6 Gaze heading signal, velocity and acceleration profile

	AT		FT		AVT	
	60°	30°	60°	30°	60°	30°
Time						
MT (s)	0.618 ± 0.347	0.072 ± 0.017	0.161 ± 0.036	0.109 ± 0.102	0.196 ± 0.133	0.123 ± 0.165
GL (s)	0.987 ± 0.492	1.026 ± 0.322	1.210 ± 0.687	0.874 ± 0.299	0.548 ± 0.199	0.638 ± 0.241
TT (s)	1.605 ± 0.181	1.099 ± 0.326	1.371 ± 0.683	0.983 ± 0.287	0.744 ± 0.096	0.761 ± 0.109
Delta						
DG (°)	52.24 ± 6.60	27.39 ± 7.08	55.76 ± 4.67	29.58 ± 5.05	58.43 ± 5.66	32.00 ± 9.09
DH (°)	27.72 ± 8.06	5.64 ± 5.25	30.24 ± 14.16	14.84 ± 9.32	27.11 ± 8.68	7.79 ± 6.02
DE (°)	24.51 ± 3.64	21.75 ± 8.36	25.52 ± 14.23	14.74 ± 10.67	31.32 ± 7.20	24.21 ± 4.66

Table 1 Mean values of the main parameters and their relative standard deviation (SD)

2.3 Results

Experimental tests gave important results about the capability of the proposed system to track the gaze during the execution of the experimental paradigms. The application of the criteria described in the data selection paragraph brought about a reduction of the total amount of data with a final data loss of approximately 40 %. Figure 7 shows a typical attention task in which an infant orients his/her attention to the periphery (left: from screen 3 to 4, i.e. 30°, right: from screen 3 to 5, i.e. 60°). More specifically, it shows how transition of the gaze from the central screen (#3) to the peripheral one (#4 or #5) works and data about gaze and head position return information about the contribution of head and eye during the required movement. The proposed system allows to track the gaze and the head in the whole AOI. A good quality of head position is present when the head is found and tracked in at least two cameras of the eye-tracker. Thanks to the presence of six cameras, the head is almost visible in two cameras during the entire duration of the tasks. This particular feature reduces the data loss and inaccuracies that might result from nonoptimal head orientations revealed in other works (30). It is worth to mention also noise issue, the SmartEye system gives the possibility of setting parameters that affects the gaze output values called 'Filtered' and consequently the signal noise (e.g. saccade and pupil filters). In details, it is possible to specify at which angle an eye movement will be classified as a saccade and how long the fixation filter should be. In addition, the diameter of the pupil does not change very rapidly; temporal filtering is therefore used successfully to reduce the amount of noise. From data analysis of the three tasks, it is possible to observe relevant information presented in the following sections about system capabilities to measure infant behavior in all of the three paradigms. Results are presented on the basis of the defined parameters.

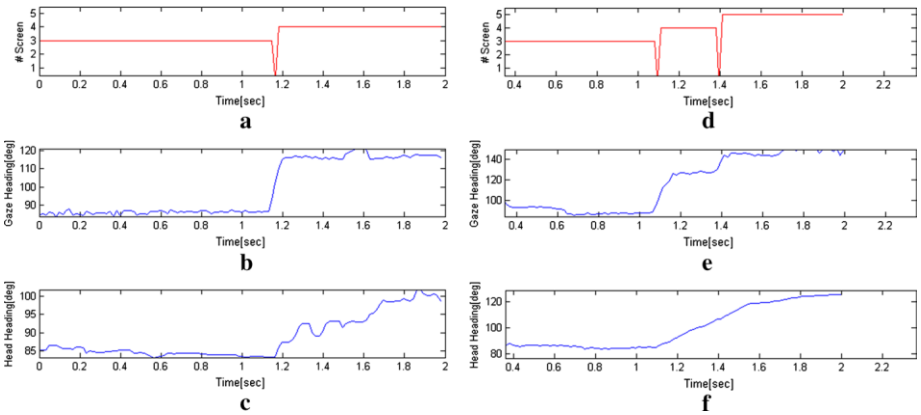


Fig. 7 Examples of attention task: a results of SmartEye analysis of intersection between the gaze with the screens. During the transition from screen #3 to #4, the gaze passes through the space between the two screens thus it does not intersect one of the AOI and the system returns zero value, b gaze heading during the transition from screen #3 to screen #4, c head heading during the transition from screen #3 to screen #4, d results of SmartEye analysis of intersection between the gaze with the peripheral screens. In this case the system returns zero value when the gaze is between screen 3 and 4 and between 4 and 5, e gaze heading during the transition from screen #3 to screen #5, f head heading during the transition from screen #3 to screen #5

Time parameters

In all the three tasks, MT values are lower when stimuli are presented at 30° with compared to 60° (AT: $p = 0.001$, FT: $p = 0.003$, AVT: $p = 0.015$) (Fig. 8). Attention tasks present similar values of GL at 30° and 60° ($p > 0.05$) as in audio-visual ones ($p > 0.05$), while in the fixation task it is possible to observe slightly higher and more variable values at 60° compared to 30°($p > 0.05$). It is worth saying that 60° values present high variability (Fig. 8).

Since TT is the sum of the previous values, we can observe that in the attention task, TT values are higher at 60° compared to 30° ($p = 0.003$). The same trend is present in the fixation task even if it is not significant ($p > 0.05$). This may be due to the high variability of the values at 60°. On the contrary, the audio-visual task presents similar values of TT at 30° and 60° ($p > 0.05$) (Fig. 8).

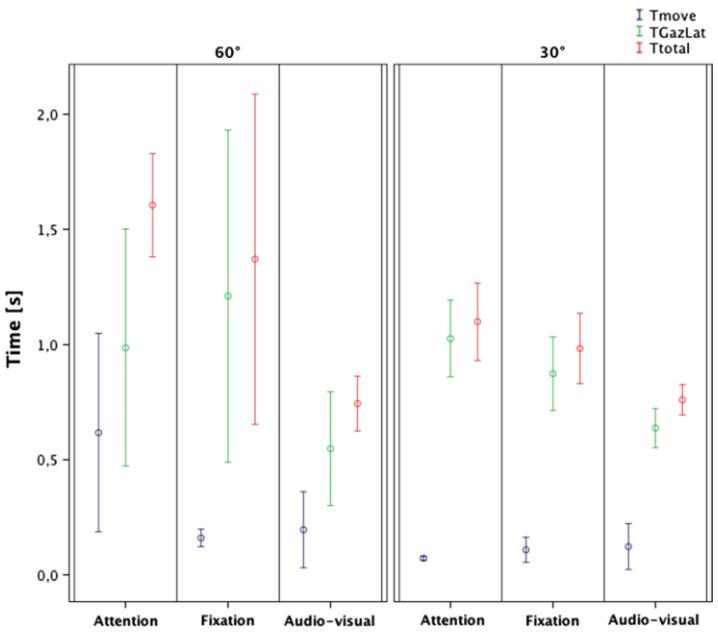


Fig. 8 Time parameters results

Comparison of time parameters between attention and audio-visual tasks

In the audio-visual task, at 30° TT and GL are largely lower than in the attention (TT: $p = 0.016$, GL: $p = 0.008$), while MT presents the same value ($p > 0.05$). At 60°, TT is largely lower in the audio-visual task ($p < 0.009$) while the other two parameters are slightly lower (MT: $p = 0.047$, GL: $p > 0.05$).

Delta parameters

By comparing values between 60° and 30°, in all three tasks, DG and DH values at 60° are significantly higher than at 30° (DG: AT $p = 0.001$, FT $p < 0.001$, AVT $p = 0.003$; DH: AT $p= 0.001$, FT $p = 0.008$, AVT $p = 0.004$). DE values are similar in attention and fixation tasks ($p > 0.05$) and also in the audio-visual even if the difference between the two values are limited ($p = 0.04$). In the attention and audio-visual tasks, at 30°, it is possible to observe that DE values are higher than DH values (AT: $p < 0.001$, AVT: $p < 0.001$) while they are similar in fixation task ($p > 0.05$). At 60°, DE values and DH values are similar in all the three tasks ($p > 0.05$) (Fig. 9).

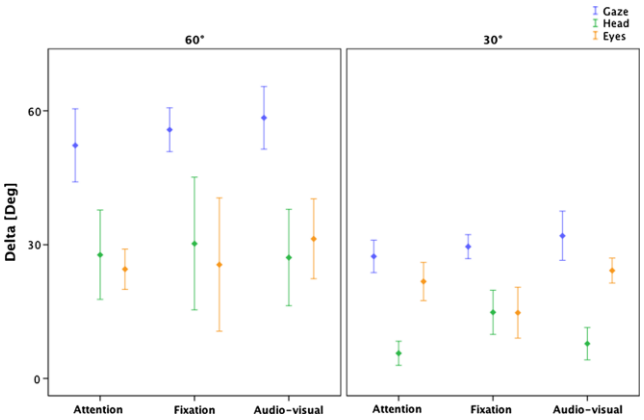


Fig. 9 delta parameters results

2.4 Discussion

In the last decade there has been an explosion of research using eye tracking with infants thanks to the evolution of the technological solution. However, automatic eye tracking presents several challenges such as the need for a good calibration procedure, the need for a purposeful experimental paradigm for infants and the difficulties of data processing. In this work, the technological challenge was to build a system able to measure an infant's gaze in a wide visual field covering a total visual range of ±60° from the centre with an intermediate evaluation at ±30°. Moreover, the same system, thanks to different integrated software, was able to provide different visual paradigms (as gap, overlap and multisensory) assessing and comparing different visual sub-competencies. The proposed system endowed the integration of a commercial eye-tracker into a purposive setup in a smart and innovative way. The calibration procedures of the system and the infant's gaze allowed us to obtain reliable data with an accuracy less than 1°. One encouraging result is represented by the feasibility of the assessment. The infants performed the three tasks and the system acquired their quantitative data. The infants maintained the sitting posture with ease and the environment allowed the infant to only pay attention to the stimuli, avoiding the distracting factors. Also the duration of the experiment was well calibrated: all the infants managed to

successfully complete all the three tasks. The results provided two kinds of data: delta and time values. The system allowed us to obtain detailed information of each task about the gaze shift and visual attention in terms of distance and movement of the head and eyes as well as duration and latency. Additionally, it was possible to make comparisons in each task between 30° and 60° and also between tasks (AT vs AUV).

Compared with previous studies, the expectations and the past findings have been obtained, and it has been possible to study the behaviour at two distances from the centre of the visual field: 30 and 60 visual degrees. Even if the small sample cannot give information and generalisations about visual functioning, some interesting findings are promising for the use of CareToy C in the clinical practice. The main strength is that the same system can measure different abilities in the same infant and across infants within different ages, exploring a wide visual field and distinguishing the eye and the head component contributions to the gaze. An interesting finding about time parameters is that the system was able to discriminate between different behaviour in presence or absence of sound stimuli, confirming the literature that the speed is higher in the presence of sound. In particular, the CareToy C system detected that with the presence of sound the TT values are similar at 30° and 60°. Comparing audio-visual (AVT) and attention (AT) tasks (it is worth underlining that those tasks are identical except for the spatial associations of the visual stimulus with the sound) the system has shown, in the AVT a faster visual response at 30° with lower values of the GL and at 60° lower values not only GL but also of MT. Another interesting result is that the system highlighted higher variability of the TT, related to the GL component in the overlap condition, probably due to the presence of a competitive stimulus (fixation task) which makes the task challenging at 60°. Regarding delta parameters, the analysis of DG confirms the reliability of the CareToy C system because, as expected, the values are in the range of 30° or 60° demonstrating that a larger movement is necessary to reach the more peripheral areas of the visual field. Moreover, the system is able to quantitatively distinguish eye and head components. An interesting finding,

to be confirmed in a larger group of infants, is that stimuli at 30° seem to be visually detected using mainly eye movements, while for larger movements (60°) a head compensatory movement is required. This behaviour is evident in attention and audio-visual tasks, instead of fixation task, in which the head contribution is already evident at 30°; this could be due to the presence of a competitive stimulus and the relative difficulty in the gaze disengagement. Further studies on the different strategies at 30° and 60° across different ages could be very interesting for new hypotheses of visual development and applications (e.g. new treatment strategies).

This study presents some limitations. First of all, the data loss remains a critical point; future work will investigate algorithms for improving data quality and the possibility to trigger the stimuli on the screens on the basis of the current position of the gaze obtained with the eye-tracker. Furthermore, the sample size of this study was quite small, but it allows us to demonstrate the feasibility of the purpose. An interesting future development could be to test a wider sample in order to obtain quantitative about the general development of visual perception in infancy.

2.5 Final remarks

This new system has been demonstrated tool for measuring and evaluating infants' gaze capabilities in a wide visual field and it is useful in providing quantitative data that can enrich the clinical assessment and in particular for objectively evaluating changes after a treatment. Although the traditional assessments are recommended, a more quantitative approach was missing in the clinical setting. The proposed approach could provide important insights on the gaze movements of very young or non-collaborative infants and could facilitate the comparison in the same subject before and after a treatment or between different samples. It should be underlined, however, that the proposed approach couldn't be separated from a global assessment of the infant, that is not exclusively related to the visual abilities, but also postural competences or social-emotional level. Our quantitative assessment is not intended then as a substitute for the traditional clinical assessment, but as a complementary tool for a quantitative analysis of gaze movements and visual field.

As a further improvement for this study could be to enhance the acquired data, reducing missing data trying in parallel to automatize and optimize the gaze calibration.

The recording of eye movements in very young infants is a real challenge and our set-up has just in part reduced the related issues, but there are still many factors that can influence the acquisition, making sometimes the trial unsuccessful. The main problem is represented by the fact that it is not easy to be sure where the infant is looking and sometimes the trial results not valid only because the infant has not the required attention in the specific moment of the presentation of the stimulus. For this reason, an important milestone for the future will be to enrich the system with the on-line control, which allows to customize the presentation of the stimuli and feedbacks directly during the execution of the tasks.

Data reported in this Chapter are based on the following paper:

Pratesi A, Cecchi F, **Beani E**, Sgandurra G, Cioni G, Laschi C, Dario P, *A new system for quantitative evaluation of infant gaze capabilities in a wide visual field* Biomed Eng Online. 2015 Sep 7;14:83. doi: 10.1186/s12938-015-0076-7.

CHAPTER 3

CARETOY H: QUANTITATIVE ASSESSMENT AND INTERVENTION APPROACHES FOR INFANTS AT RISK FOR NEURODEVELOPMENTAL DISORDERS

3.1 Introduction

Neurodevelopmental disorders are impairments of brain growth and development affecting several brain functions including cognitive, motor, language, learning, and behavioural disorders due to many causes – genetic, lesional and environmental. Infants at high risk for neurodevelopmental disorders can be identified early, i.e. in the first weeks or months of life, through careful clinical evaluation (i.e. developmental tests, neurological examination, observation of spontaneous movement patterns) combined with specific technical tools such as neuroimaging (cranial ultrasounds, brain magnetic resonance imaging (MRI)), neurophysiological tests (e.g. electroencephalography, evoked potentials), and genetic tests (karyotype, comparative genomic hybridization-microarray). As seen before, the application of evidence-based recommendations or decision-making processes, which combine the use of clinical and technical tools at a proper point during development, is crucial for early detection of infants at risk for neurodevelopmental disorders by clinicians.

An example of early identification of infants at risk for cerebral palsy (CP), one of the most common neurodevelopmental disorders, is the combined use of Prechtl's General Movement Assessment and brain MRI. In high-risk infants these techniques display high sensitivity and specificity starting from the first months of life (i.e. General Movement Assessment, 98% and 91% respectively; MRI performed at term 86–100% and 89–97% respectively), and therefore are important tools for predicting CP. Prematurity is a high risk factor for long-term visual, cognitive and psychosocial impairments. Evidence suggests a direct causative link of poor visual function, motor delay, neurodevelopmental disorders with prematurity and prolonged recovery in Neonatal Intensive Care Unit (NICU). As stated by the World Health Organization, identification of the infant at risk for a neurodevelopmental disorder is a crucial starting point to establish a close relationship between

parents and health care providers and to provide early intervention. The goal of early intervention is to prevent or minimize motor, cognitive, emotional impairments in young children disadvantaged by biological or environmental risk factors. Evidence suggests that brain development should be considered the result of the complex interaction between genes, social and physical environment thanks to the modulation of gene transcription and expression (epigenetic mechanisms).

The importance of the early identification and diagnosis, and above all early intervention, is highly connected to brain plasticity, that is capacity of the central nervous system to modify its structure and function. In the normal brain it is possible to distinguish three types of plasticity: experience-expectant, experience-independent and experience-dependent. The complexity and sensitivity of the neonatal brain to environmental stimuli is referred to as experience-dependent plasticity, i.e. the modifications of structural and functional brain pathways in response to the 'rendezvous' between genes, environmental stimuli, and experiences. Even if this plasticity occurs throughout life, young brains show greater potentiality towards this phenomenon, thanks to prominent mechanisms of myelinization, creation, and sprouting of neural projections essential for brain development and modelling of cortical neuronal circuitries. Critical periods of brain plasticity are defined as those periods in which development of brain functional properties are strongly dependent and shaped by experience and environmental stimuli. Evidence of critical periods for visual, auditory, somatosensory systems, and cognitive functions has been largely demonstrated.

Early intervention should begin when infants are still in the Neonatal Intensive Care Unit, mainly by focusing on reduction/minimization of stress factors, or soon after Neonatal Intensive Care Unit discharge. The main aim of early intervention programs after hospital discharge is no longer the reduction of stress but to promote infant development. They are focused on parent–infant relationship, infant development, or both, with the aim of improving the overall functional outcome of preterm infants. However, high heterogeneity among

early intervention programs (e.g. varied background and theoretical constructs, timing, duration, etc.) and lack of a high quality randomized controlled trial studies limit conclusions.

After the discharge from hospital, the intervention should move to an environmental enrichment, that could be defines as the adaptation of infant's environment, that is the stimulation of the brain by its physical and social surrounding. Strong evidence of positive effects of enriched environment on brain plasticity, both in physiological and pathological conditions, has been suggested. Enriched environment could represent a potential 'behavioural therapy'. Recently a definition for enriched environment interventions in infants has been proposed by Morgan et al. that describes enriched environment as 'interventions aimed to enrich at least one of the motor, cognitive, sensory, or social aspects of the infant's environment for the purposes of promoting learning'. Open issues questioning the effectiveness of early intervention programs are heterogeneity of intensity, focus, setting, and participants in studies. However, clinical and experimental findings seem to indicate that, to be maximally effective, early intervention has to be early, intensive, active, tailored for each individual and family-centred. Intervention programs that fully satisfy these essential criteria are expensive, although economically justifiable if compared with costs of long-term disability. The possibility of creating a common family-centred home setting, which is able to quantitatively measure activities and progress and be remotely managed by rehabilitation staff, can represent a good solution. Therefore, biotechnologies and tele-rehabilitation could represent a promising approach to provide home early intervention programs for a large number of infants, at a relatively low cost. In this field, neurodevelopmental engineering is a new ground-breaking interdisciplinary area, at the crossroads of developmental neuroscience and bioengineering, aiming to provide new methods and tools for quantitative analysis and modelling of human behaviour during typical and atypical neurodevelopment. It is mainly devoted to developing new clinical protocols and standards to help clinicians perform early diagnoses, functional

evaluations, and interventions of neurodevelopmental disorders by providing new generations of educational and interactive toys that can provide quantitative measures, adequate stimuli, and support for psychomotor development. A variety of sensorized systems, such as toys able to measure grasp-force and grasping-shape, and other technology-assisted methods to obtain early diagnostic information for neurological disorders, are summarized in a recent review. Moreover, these new technologies have been proposed also for home rehabilitation. A recent European research project, called CareToy (www.caretoy.eu), aimed to clinically evaluate an intensive, individualized, home-based, family-centred early intervention programs and consisted in a new technological smart modular system, remotely monitored by a rehabilitation staff. These new frontiers in rehabilitation are completely in line with the guidelines of Convention on the Rights of Persons with Disabilities regarding the requirements of assistive technology and Information and communication technologies and the need to promote information and communication technologies policies and programs in the field of rehabilitation, as stated in a recent research paper. In this framework, the biomechatronic techniques, that is, techniques integrating aspects of biology, mechanics and electronics, can open the possibility of managing early intervention in infants with neurodevelopmental disorders at home, far from the clinical center, expanding the access of infants to early intervention in the near future.Early detection and diagnosis of neurodevelopmental disorders is crucial in order to start early intervention, as soon as possible, optimize the use of limited resources for infants who need these resources most, to positively modify natural history of disorders. Early diagnosis is possible starting from the first months of life thanks to a careful clinical evaluation,which combines the use of genetic and neuroimaging tools. For this reason, common efforts need to be directed at identifying evidence-based recommendations for diagnosis and management of neurodevelopmental disorders. Moreover, as confirmed by clinical studies, parents should be involved in both processes (early diagnosis and intervention).

3.2 Sensorized toys for measuring manipulation capabilities of infants at home

Preterm infants, i.e. babies born after a gestation period shorter than 37 weeks, spend less time exploring objects. The quantitative measurement of grasping actions and forces in infants can give insights on their typical or atypical motor development. The aim of this work was to test a new tool, a kit of sensorized toys, to longitudinally measure, monitor and promote preterm infants' manipulation capabilities with a purposive training in an ecological environment. This study presents preliminary analysis of grasping activity. Three preterm infants performed 4 weeks of daily training at home. Sensorized toys with embedded pressure sensors were used as part of the training to allow quantitative analysis of grasping (pressure and acceleration applied to toys while playing). Each toy was placed on the midline, while the infant was in supine position. Preliminary data show differences in the grasping parameters in relation to infants age and the performed daily training.

3.2.1 Materials and methods

The analysis of manipulation capabilities in infants has been possible thanks to a purposive experimental setup, carefully studied to focus the infants attention on specific targets. Fostering the subject to grasp a sensorized object allowed us to measure a set of quantities related to his/her interactivity with the device, which depends on the responsiveness to provided stimuli and development of motor and grasping skills. Such a study has required a reproducible test environment, to allow for evaluation of a subject's progresses during experiment time, and for comparison among the performances of different subjects. Due to the noncollaborative nature of infants, this environment had to be as familiar as possible to them, in order not to alter their spontaneous behavior, allowing us to evaluate their manipulation capabilities in normal

conditions. Still, as said, the environment must provide stimuli capable of inducing the handling of sensitive objects, which we will call toys from now on. Experimentations have consisted in so-called training sessions, in which the toys were presented to the infant. Data collected from the sensors during training sessions have been analyzed with numerical tools and techniques to extract the manipulation parameters of interest. A more detailed description of the main tools and techniques used follows.

Experimental setup

The experimental setup is composed of: i) a kit of sensorized toys, ii) an arch, iii) a laptop. Based on the concept of affordance a kit of 3 toys, inspired by existing toys realized for infants, has been designed and developed. A purposive electronic board has been developed around the iNemo by ST Microelectronics® . Toys hardware counts: pressure sensors (LPS331AP), inertial sensor (iNemo), force sensors (FSR by Interlink) and multicolour LEDs for light feedback. While the iNemo is inserted in the case of the toy, the other two sensors (force and pressure) are embedded into silicone and compose the sensitive part. Toys circuitry includes the components for interfacing the two sensors embedded in the silicone lighting the LEDs, while stable connection with a laptop is guaranteed by a pushpull aluminium connector. Toys are hung over the baby on their cable from a purposive arch, in order to encourage the manipulation during the training. Signals are acquired via USB on a laptop and saved for off-line analysis. Three basic shapes have been chosen for the sensorized toys, on the basis of the type of grasp (palmar or digital) and hand action (uni or bimanual) that we want to detect and monitor. The shapes have been carefully chosen to have a clear affordance and encourage different manipulation approaches. The mouse toy is a cylinder equipped with one pressure sensor and lights inside the main body of the toy to attract infants' attention, with two small petals embedding a force sensor and a light in their central part. The U-toy is a horseshoe composed of two cylinders with a pressure sensor and lights in each side. Eventually, the ring toy is a ring

equipped with pressure sensors and lights inside the sensitive part. The final toys are toys shown in Figure 1. Thanks to this easy setup, the sensorized toys can be directly used at infants' home.

Fig. 1: sensorized toys: Ring, Mouse and U toy

Training session

This study comprised three preterm infants that met the following criteria: i) gestational age 28 + 0 weeks and 32+6 weeks; ii) corrected age (CA) at baseline between 3 and 9 months; iii) appropriate for milestone in gross-motor abilities. Infants with brain damage i.e. brain malformation, severe sensory deficits (blindness, deafness) or other severe non neurological malformations, presence of intra-ventricular haemorrhage (IVH) > 1, any degree of periventricular leukomalacia (PVL), known epilepsy or other form of seizure were excluded. The clinical protocol is deeply described in (13). Three preterm infants (3, 4 and 5 months old of CA) performed daily 4 weeks of training at home. The sensorized toys were used as part of the training to obtain quantitative analysis of grasping. A toy was placed on the midline while the infant was in supine position. The lights inside the sensorized toys blinked in order to encourage the grasping and the infant freely manipulated the toy hung on the arch (Figure 2). Each training session lasted about three minutes for a total training time of 1h 45m (std: 38m). The study has been conducted in Italy and approved by the Ethics Committee of Pisa University Hospital (Italy) and Tuscan Region Pediatric Ethics Committee (Italy).

Fig. 2 An example of traning session: the infant is manipulating the sensorized ring toy on the arch

Manipulation parameters definition

In this analysis we considered the signals coming from the pressure sensors (thus generated by palmar grasps), the accelerometer and the magnetometer embedded in each toy. Retrieved signals are analyzed independently from the toy model and exploited to compute four parameters quantifying the intensity of manipulative activity during a session:

Mean grasp length: Expressed in seconds (s), is the mean interval between starting and ending time of a single grasping action. Grasps are assumed to start when a pressure value greater than 100 Pa (triggering threshold) is continuously detected for more than one second. The end of grasps is taken as the instant after which the pressure signal keeps below the threshold.

Mean grasp intensity: Expressed in pascal (Pa), is the mean value acquired by the toy's pressure transducers during a grasp.

Grasping time over training time: Dimensionless, it is the mean value of the grasping time over the total training time.

Mean acceleration: Expressed in m/s2 , is the mean value of acceleration norm during grasping intervals. We assume that this parameter is proportional to the force applied by the infant to the toy to move while playing. The above

parameters have been extracted from raw data writing MATLAB® scripts compatible with a release R2013a basic installation, with no need of additional toolboxes.

3.2.2 Results and discussion

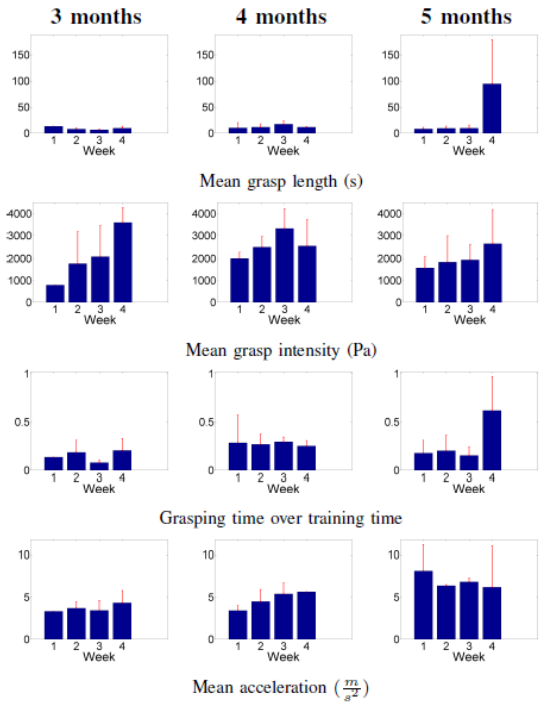


Fig. 4 Manipulation parameters displayed for the tree infants

Figure 4 shows the four parameters computed for the three infants. Data across the trainings show a positive trend in all the 3 infants, with an increased mean grasp intensity. This finding is consistent with previous data on healthy term infants adding the possibility that the training can accelerate this trend. Moreover, other considerations are possible by analyzing the trend of each infant. The younger infant (3 months old) shows a high increase of grasp intensity, a slightly increase of the acceleration while weeks. These findings could be related to the small age of the infant that modified during the training the grasp intensity while grasping with a related increase of acceleration. As far as the 4 months old infants is concerned, we can observe an increase not

only in the grasp intensity but also in the mean acceleration. The last one finding could be related to a higher manipulation activity during the grasps. The 5 months old infant performs more grasping actions, also increasing intensity. In the fourth week of training there is an evident increase of grasps duration with a decrease of acceleration. From a clinical point of view, this finding could be related to an increased manipulation activity associated with smoother and more movements that reduce the acceleration. There are some methodological limitations of the current study that are important to highlight. The small sample size examined and the preliminary data analysis and techniques are evident limits of the current study. Sensor-fusion of gyroscope and magnetometer data could improve the immunity of acceleration retrieval to variable magnetic interferences, and will be introduced as one of the most urgent refinements to data analysis method. Further analysis will focus on additional manipulation parameters depending on the different shapes and relative affordance of the toys.

3.3 CareToy: stimulation and assessment of preterm infant's activity using a novel sensorized system

Early intervention programs aim at improving cognitive and motor outcomes of preterm infants. Intensive custom-tailored training activities are usually accompanied by assessment procedures, which have shortcomings, such as subjectivity, complex setups, and need for structured environments. A novel sensorized system, called CareToy, was designed to provide stimulation in the form of goal-directed activity training scenarios and motor pattern assessment of main developmental milestones, such as rolling activity, grasping, and postural stability. A group of 28 differently skilled preterm infants were enrolled. Acquired measurement data were analyzed with dedicated sensor data processing algorithms, along with clinical evaluation of motor ability. High correlation among technically determined parameters and Alberta Infant Motor Scale values was determined by Pearson correlation coefficients. Due to good accuracy and possibility of single motor skill subfield analysis, results confirm system suitability for motor ability assessment. Statistical analysis of inter-motor ability group and intertraining goal data comparisons demonstrate system's appropriateness for goal-directed activity stimulation. CareToy has evident potential of being an important contribution to the field of infant motor development assessment, expanding accessibility of early intervention programs and affecting rehabilitation effectiveness of preterm infants.

3.3.1 Materials and methods

This section first presents characteristics of the CareToy platform, followed by description of the proposed outcome measures, including the corresponding posture, movement, grasping, and stability assessment algorithms, as well as applied clinical assessment scales. Afterwards, details of enrolled infants are given. Finally, a description of methods for statistical analysis is provided.

CareToy System

The CareToy platform, designed for the home environment, is composed of different modules (Fig. 1), aiming to stimulate and measure infant's actions. The system is infant friendly and has during the process of prototype development also passed the tests About conformity to safety requirements for a medical device with a CE mark. The instrumented baby gym is equipped with interactive walls for audio-visual stimulation of infant attention, activity, and gaze movements. This is achieved with a monitor in the frontal wall for showing short videos, along with coloured lights, speakers, and switches in the lateral walls. Floor is covered with two Tekscan pressure mattresses (CONFORMat System, model: 5330) for pressure distribution and infant posture assessment. Each pressure mattress has 1024 force-resistive sensors, whereas pressure data are sampled with 30 Hz. Wearable wireless magneto-inertial measurement units (IMU) on trunk and forearms are intended for postural control measurement, as is in full detail also described in Rihar et al. (2014) Battery powered IMUs that comprise three-dimensional gyroscopes (L3GD20), accelerometers and magnetometers (LSM303DLHC) were developed by STMicroelectronics and are integrated in specially designed bracelets and chest strap. These bracelets have additional dedicated artificial silicone soft covers of neutral beige colour. Along with the light weight and minimum size of IMUs this en sures negligible effect on characteristics of infants' behavior. Moreover, IMUs were already placed on infant's wrists and trunk some minutes before the beginning of training, which helped the infants to become familiar with IMUs, also reducing the behavioral effects, related to having something new on the body. Sampling frequency of IMUs is 100 Hz. sures negligible effect on characteristics of infants' behavior. Moreover, IMUs were already placed on infant's wrists and trunk some minutes before the beginning of training, which helped the infants to become familiar with IMUs, also reducing the behavioral effects, related to having something new on the body.

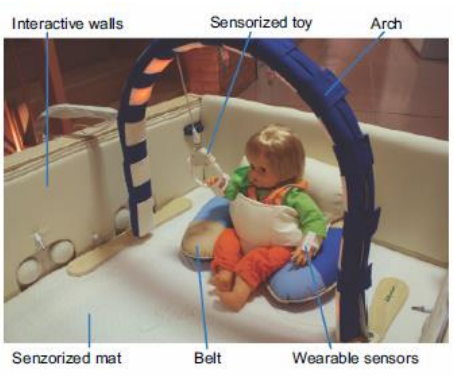


Fig. 1 CareToy platform and its main modules

Sampling frequency of IMUs is 100 Hz. A central arch with 12 lights and connectors for mechatronic toys34 serves for gaze movement stimulation and manipulation capability assessment. The purposively developed toys were inspired by commercial toys for infants and were designed on the basis of affordance to encourage different manipulation approaches and to be compliant with infants' hand dimensions. Different basic shapes and sizes, including toroidal and cylindrical shape ensure grasping variability. Toys include multicolour LEDs for stimulation and light feedback, along with integrated pressure sensors, force sensors, and IMUs for toy interaction assessment. Sampling frequency of sensors in toys is 100 Hz. The tele-rehabilitation module allows the system to remotely communicate with the clinical staff for monitoring and assessing the rehabilitation techniques. Data is stored on a server for postprocessing. CareToy Clinical represents another part of the larger CareToy environment but is not directly relevant for the present study. It was designed to be located in the clinical centres and comprises five video screens and an eye tracker for stimulation and analysis of infant gaze characteristics. Full details of CareToy Clinical are given in chapter 2.

Outcome Measures

This subsection provides a description of the selected outcome measures and the corresponding applied algorithms for sensor data extraction, processing, and analysis, consecutively presenting the various addressed fields of infant

activity. Rolling Range-of-Motion (ROM) IMU sensor data, namely vectors of angular velocity, acceleration, and magnetic field are for each performed scenario collected throughout the entire training session with a sampling frequency of 100 Hz. Data are first processed by applying the unscented Kalman filter (UKF), which was implemented similar as in Beravs et al. No calibration step is needed to orient the IMUs, as consistent placement is achieved by proper labelling of IMU sensor covers. The evaluated trunk IMU orientation is then expressed relative to the gym IMU coordinate system. To minimalize the possible effects of potential trunk IMU displacement, trunk IMU orientation data are corrected by taking into account the trunk orientation, estimated from the pressure imprint data. The full sensory data fusion procedure is in detail given in Rihar et al. The range of rolling angle from supine to prone and back to supine is 360°. By using the improved trunk orientation information, the rolling range-of-motion (ROM) parameter is calculated for each measurement session is the angular distance between the values of 90th and 10th percentile of the angular rolling data.

Forearm Orientation Intensity and Area

Forearm (FA) IMU sensor data are also processed with the UKF and expressed in the trunk coordinate system to determine the FA orientation relative to the trunk posture. Afterwards, orientation data are recalculated into two orientation angles, namely elevation and azimuth. Elevation is calculated as the angle between the FA orientation vector and the coronal plane, while the azimuth is determined as the angle between the sagittal plane and the FA orientation vector projection onto the coronal plane. Both angles describe FA orientation in a sphere, whereas range of azimuth is 360° and range of elevation angle is 180° (see Fig. 2). Following this, spherical orientation data is transformed into planar presentation, whereas orientation angle data (azimuth and elevation) are first grouped into small areas of 3° by 3° to reduce data amount, and further organized with azimuth and elevation data on horizontal and vertical axis, respectively. This produces a FA orientation map for each arm (see Fig. 2).

Following this, FA orientation intensity parameter is calculated as the percentage of session duration with the FA oriented in lateral-medial direction. Lateralmedial FA orientation area parameter is determined as the percentage of frontal lateral-medial orientation map subpart area, covered with the FA orientation data. Finally, parameter values are averaged for both arms.

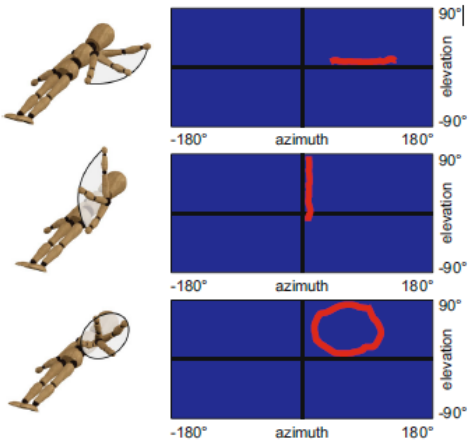


Fig. 2 Representation of the FA orientation map determination for arm movements with changes of azimuth angle (top subplot), elevation angle (middle subplot), and both angles simultaneously (bottom subplot).

Toy Grasping and Toy-Hand Interaction Percentage

Toy grasping and toy-hand interaction actions are analyzed by processing data of force, pressure, and IMU sensors, integrated in the mechatronic toys34 (see ‘‘CareToy System’’ section). Grasping activity in this context covers all toy grasping activities with sufficient grasp strength. First low-pass filtering and signal trend removal methods are applied on the pressure sensor signal to remove high-frequency noise and signal drifting character, which can be a consequence of the pressure chamber air temperature changes. Afterwards, threshold comparison and data connectivity method are used on the pressure signal to extract and determine the pressure-based grasping activity intervals with adequate grasp strength. Potential artefacts, such as extremely short grasping intervals are removed from further analysis. Maximum and mean

values of the pressure signal are calculated. Since force sensor signal is less subject to drift, only low-pass filtering method is applied along with threshold comparison and signal data grouping to extract the force sensor-based grasping activity intervals. Unity of both signals (pressure and force) is used to assess the general grasping activity. Toy grasping percentage is determined as the ratio between grasping time and session duration for each measurement session. Toy-hand interaction activity in this context covers all activities, when infant was in contact with the toy, herewith included grasping the toy, touching the toy without grasping, and hitting the toy. With this in mind, interaction starts as soon as infant touches the toy and ends when infant loses contact with the toy. To evaluate interaction activity, toy IMU gyroscope data are first transformed to the gym coordinate system. Low-pass filtering and wavelet transform are applied on the angular velocity data to determine the toy oscillation intervals, which ensure toy movement without interaction. Following this, signal energy of angular velocity data is calculated and compared to an empirically pre-determined threshold to determine the toy movement intervals. This threshold represents the maximum signal energy of angular velocity data of mechatronic toys in standstill. Removal of toy oscillation intervals from the toy movement intervals determines the interaction intervals. Interaction percentage is determined as the ratio between the acquired interaction time and session duration.

COP Movement Parameters

The characteristics of centre-of-pressure (COP) movement of infants in sitting position are used to evaluate the level of infant's postural stability.13,20 COP movement is assessed based on the pressure imprint data, which are sampled with 30 Hz. First, pressure imprint data are pre-processed with bias and superposed noise removal methods. Following this, the buttocks pressure imprint is extracted by taking into account the pre-determined position settings and thus removing potential imprints of the dedicated belt pillow. Centre-of-pressure coordinates of the buttocks pressure imprint are calculated for the

entire measurement session. The obtained COP movement vector is then for each training trial analysed by calculation of some well-established parameters, such as root-meansquare displacement, circle area, along with anteriorposterior (AP) and medial–lateral (ML) range-of-motion.

Alberta Infant Motor Scale

The infants were assessed with a battery of clinical tests41 before and after training with the CareToy system. For the current work, the main clinical test was the Alberta Infant Motor Scale (AIMS). This standardized scale, used in infants from term until 18 months of age, assesses infant's motor abilities, quality of posture and movement outcomes in four positions: prone, supine, sitting, and standing. It is a reliable measurement in detecting delayed and abnormal motor development.

Subjects

According to the clinical study protocol, eligible infants were preterm infants born between 28 + 0 and 32 + 6 (weeks + days) of gestational age, aged 3–9 months (12–38 weeks) of corrected age (CA), who had achieved a predefined cut-off score in gross motor ability, derived from Ages & Stages Questionnaire Third Edition (ASQ-3). Exclusion Criteria were: (i) birth weight below the 10th percentile (infants small for gestational age); (ii) brain damage i.e. intra-ventricular haemorrhage (IVH) more than grade 1, any degree of periventricular leukomalacia (PVL), or brain malformation; (iii) epilepsy or other form of seizure; (iv) severe sensory deficits (blindness, deafness) and v) other severe nonneurological malformations. These were determined at the onset of the CareToy project during definition of the clinical study protocol by the clinical staff (child neurologists and neonatologists). At the beginning of the enrolment phase, they were evaluated by screening the clinical history of infants, admitted at the Neonatal Intensive Care Unit (NICU) of Santa Chiara Hospital in Pisa, Italy. 28 preterm infants with CA between 13 and 27 weeks and with AIMS score between 8 and 23 at the beginning of the training were enrolled. The infants were equally distributed into three groups, based on their

AIMS scores: 9 were included in the lower group (AIMS £ 10), 12 in the middle group (10<AIMS £ 16), and 7 in the higher group (16<AIMS). Such cutoff criteria were chosen to enable similar differences among group motor skill levels. The study has been approved by the Ethics Committee of Pisa University Hospital (Italy) and Tuscan Region Paediatric Ethics Committee (Italy), and it has involved the clinical centre IRCCS Fondazione Stella Maris, Department of Developmental Neuroscience, in Pisa (Italy), in collaboration with Neonatal Intensive Care Unit, Pisa University Hospital ''Santa Chiara''. The trial is registered at ClinicalTrials.gov (NCT01990183) and adheres to the Declaration of Helsinki. All the parents of the enrolled infants provided their written informed consent to enter the study.

Training Scenarios and Goals

The rehabilitation staff, including child therapists and child neurologists, selected a set of response specific goals, most important for the development course of pre-term infants during their first year of life. These among other cover promotion of rolling activity, arm movement, toy play (touching and various grasping abilities), and postural stability. Each of these goals can be promoted in various ways, therefore clinicians have designed a library of goal directed activities, namely scenarios to be promoted inside the CareToy system. These are posture-specific and can be presented to the infant, while in supine, prone or sitting position. From a technical point of view, the scenarios comprise an activation sequence of specific CareToy modules, such as LED lights, speakers, and video screen. Toys can be used in infants with higher motor abilities for promotion of reaching and grasping behaviors, but also in younger infants for promotion of attention and pre-reaching movements. Duration of each scenario is determined by the rehabilitation staff before training and is normally set between 2 min and 10 min, according to the level of complexity, training goals, as well as infant's age and general motor skill capabilities. The same training goal can be addressed by different scenarios of different lengths, whereas achieved effects can be enlarged according to their

duration. Stimuli activation can be automatic or reward- based, according to specific prefixed thresholds, such as strength of grasp or successfully accomplished rolling activities. The form of scenarios can thus be changed on the basis of infant's activities, capabilities, and success rate. High success rate and quick responses can result in quick reward-based stimuli activation and consequential shortening of training scenario's duration. This adaptation is needed and implemented to ensure that the training sequences are challenging and motivating enough, regardless of the motor skill capabilities. In case of very high success rate, rehabilitation staff can take the response-related results into account and adjust the scenario complexity before the beginning of next training. The library of scenarios is organized by each infant and position, while the scenarios are grouped by CareToy modules involved, main rehabilitation goals, and sub-goals. Although the final aim of the CareToy system is common to all infants and is focused on general progress in cognitive and motor skill development, different infants' behaviors and activities are expected in relation to their motor abilities. In this work we focused on evaluating the capability of CareToy to stimulate and assess infants' activity in supine and sitting positions. For supine position, the designed scenarios can be grouped into training goals, as follows.

Reaching and grasping toys on midline (Toy on midline).

Scenarios with this goal aim to address one of the main activities to be promoted for infant's development, namely capability of developing reach-to-grasp maneuvers. With this in mind, toys are hung in the central position of the arch for promotion of reaching and grasping activities on midline.

Reaching and grasping toys on sides of the arch, while rolling for small ranges (Toy on arch). Infants are stimulated to slightly turn their body in order to reach and grasp with ipsilateral and contralateral upper limbs the toys that are placed either on the right or left side of the arch.

Reaching and grasping toys on feedback walls, while rolling for high ranges (Toy on sidewalls). Infants are stimulated to roll in order to reach the side

position (either on right or left side of the CareToy system) and maintain this side position during toy play (reaching and grasping) with toys on lateral walls.

Head rotation (Without toys) and rolling (Rolling stimulation). The activities for this goal are mainly aimed to stimulate infants to follow with their gaze and/or with rolling of their bodies the visual stimuli, which is provided by lights on the arch and/or on lateral walls. Infant activity for the supine position related goals with toys is assessed with arm posture-based (forearm orientation intensity and area), rolling activity-based (rolling ROM), as well as toy play-based (grasping and interaction time) outcome measures. For the goals in supine position without toys, activity is analyzed by calculation of rolling ROM, as well as forearm orientation intensity and area parameters. For sitting position, scenarios can be grouped into two main goals, as follows.

Head rotation and trunk control (Without toys). Infants are stimulated by catching their visual attention (on the screen wall and/or on the feedback walls) and controlling movement of trunk and the head (maintaining head on midline, rotating head to follow visual stimuli and/or sounds).

Head rotation and trunk control while reaching and grasping toys (With toys). Infants are stimulated to reach and grasp the toys, which are placed on the arch, while maintaining head and trunk control in the sitting position. Selected outcome measures for the sitting position related goals are focused on postural stability assessment and are obtained by COP movement analysis in view of calculating parameters, such as RMSd, circle area, and AP and ML ranges. Infants were presented with numerous scenarios of all these goals during the course of 1 month training, which altogether consisted of 100–150 performed training sessions for each infant. Each infant performed up to ten mostly different scenarios per day, summing up to one and a half hours of training per day. For the purposes of this study, however only data of first five sessions for each goal were taken into account. These were evenly distributed over a few days in the first week of training. We hypothesized that the approach of analyzing data from the first week of training guarantees a reliable quantification of the variability of infants' behavior and ensures sufficient

strength (sufficiently high number of training scenarios) for performing statistical analyses, while reducing the effects of training-related behavioral changes and progress, expected throughout the course of 1 month training with the CareToy system. As duration of training scenarios is dependent on infants' success rate and motor ability, meaning that different infants probably performed different duration of training, the mean and standard deviation values of first five performed scenarios of all infants for each goal were calculated. In particular, for the supine position, Head Rotation and Rolling was assessed for 317 ± 87 s, Reaching and Grasping toys on midline (Toy on midline) for 327 ± 75 s, Reaching and Grasping toys on arch side while rolling for small ranges (Toy on arch) for 342 ± 59 s, and Reaching and Grasping toys on feedback walls, while rolling for high ranges (Toy on sidewalls) for 325 ± 62 s. For the sitting position, Head rotation and Trunk control was assessed for 225 ± 57 s and Head rotation and Trunk control, while reaching and grasping toys for 307 ± 100 s.

Statistical Analysis

Statistical analysis was applied on the large amount of data not only to provide the possibility of intergroup, inter-goal, and other comparisons, but also to ensure an intuitive interpretation. With this in mind only data of first five training sessions for each infant were used for statistical analysis of rolling, forearm orientation, interaction, and grasping based parameters. Data reduction was performed to reduce the effects of 1 month long training on the statistical measures, but still ensure a large enough amount of data to retain statistical reliability. Only COP movement data were statistically analyzed in full, because infants have performed less scenarios in sitting, thus affecting statistical measures negligibly. Kruskal–Wallis test was selected for inter-goal, as well as inter-group data comparison, providing statistical assessment of data similarity. The calculated statistically significant ($p<0.05$) and very significant ($p<0.01$) values were additionally marked with * and **, respectively. Inter-parameter correlation (for example rolling ROM and grasping percentage) and correlation among technically evaluated parameters and clinical motor

assessment scales (for example rolling ROM and AIMS scores) were estimated by calculation of Pearson correlation coefficients R. Median values of first five training sessions' data were used for the correlation calculation. Inter-parameter correlation describes the similarity of different technically determined data trends, while the second correlation coefficient denotes the similarity among technically and clinically evaluated motor skill assessment. Correlation coefficients were only determined for data pairs that were presumed relevant and clinically meaningful.

3.3.2 Results

In the following section the results of infant activity and behavior assessment are presented. First the rolling activity analysis is presented, which is followed by the results of forearm orientation intensity and area assessment. Afterwards, results on toy interaction and grasping activity are given. Finally, results of COP movement assessment parameters of infants in sitting position are provided. All results are displayed in the form of boxplots, where black horizontal lines denote the median values, the edges of the coloured boxes present the 25th and 75th percentile of data, while the whiskers extend to the most extreme data points not considered outliers. Infants are grouped in three groups according to their motor abilities by taking into account the AIMS values. Different training scenario based goals are marked with different colours.

Rolling Range-of-Motion

Rolling activity analysis results are presented with rolling ROM parameter values for four different goals (see Fig. 3), namely goals with toys on midline, arch, sidewalls, and for rolling stimulation. Statistically significant differences can be identified in case of intergoal comparison among rolling stimulation and toys on midline goals for the middle and higher groups. Inter-group

comparison demonstrates statistically significant differences especially among the middle and higher groups, as well as towards the group with lower abilities (see Fig. 3). Correlation was tested on rolling ROM data for the rolling stimulation goal and the corresponding AIMS values and is statistically very significant ($p<0.01^{**}$) with Pearson correlation coefficient of 0.71.

Forearm Orientation Intensity and Area

Activity and behavior of infant arm movement in terms of FA orientation intensity and area are first presented for goals with toys on midline and without toys (see Fig. 4), followed by results for goals with toys on midline, arch, and sidewalls (see Fig. 5). Results demonstrate statistically significant inter-goal differences, related to presence of toys (see Fig. 4), as well as statistically significant inter-goal differences, related to position of toys (see Fig. 5), especially in the lower and medium groups. Additionally, inter-group comparison reveals statistically significant differences especially when comparing outcomes towards the lower group (see Fig. 5). Correlation was tested on FA orientation intensity data for the goal with toys on arch and corresponding AIMS values, and is statistically significant ($p<0.05^{*}$) with Pearson correlation coefficient of 0.44. Pearson coefficients R for rolling ROM and FA orientation intensity for goals with toys on sidewalls and arch are 0.68 ($p<0.01^{**}$) and 0.54 ($p<0.01^{**}$), respectively. Correlation values for rolling ROM and FA orientation lateral-medial area for the same goals are 0.64 ($p<0.01^{**}$) and 0.65 ($p<0.01^{**}$).

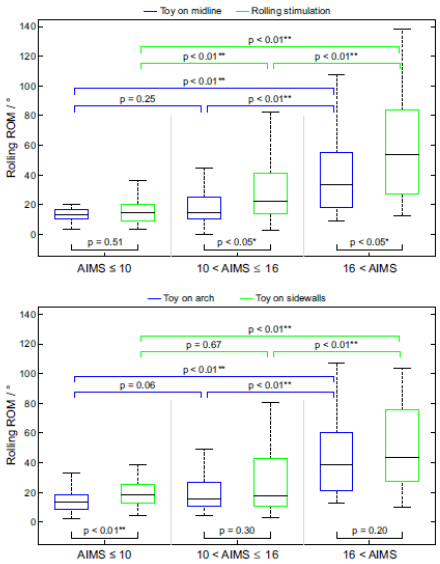


FIGURE 3. Rolling range-of-motion (ROM) parameter values for comparison of toys on midline and rolling promotion goals (upper subplot) and goals with toys on arch and sidewalls (lower subplot). Kruskal–Wallis test based *p* values are given above the coloured and black horizontal lines and describe the inter-group and inter-goal data comparison statistics, respectively.

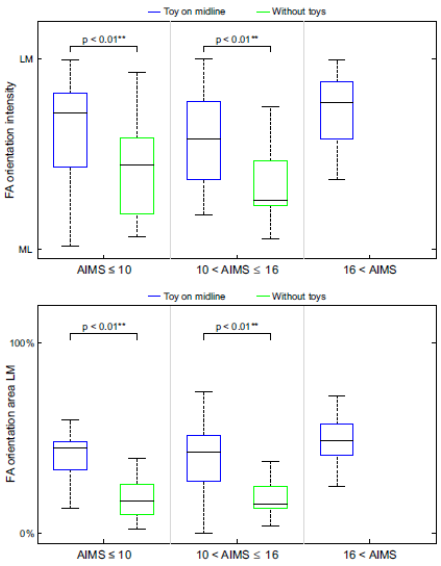


FIGURE 4. Forearm (FA) orientation intensity (upper subplot) and lateral-medial orientation area (lower subplot) values for goals with toys on midline (blue boxes) and without toys (green boxes). *p* values are given for the inter-goal data comparison. LM and ML denote lateral-medial and medial–lateral orientation.

Toy Grasping and Toy-Hand Interaction Time

Results for hand-toy interaction and grasping activity are given for three goals with toys in different positions, namely midline, arch, and sidewalls (see Fig. 6). Kruskal–Wallis test results are presented for inter-goal and inter-group comparison. Results reveal statistically significant differences for inter-goal comparison and especially for inter-group comparison, when comparing the middle and higher groups (see Fig. 6). Correlation coefficient R of interaction percentage data for midline and sidewalls, and corresponding AIMS values are 0.52 ($p<0.01$**) and 0.54 ($p<0.01$**), respectively. R values for the max grasp pressure data and AIMS values were for goals with toys on midline 0.60 ($p<0.01$**), arch 0.52 ($p<0.01$**), and sidewalls 0.62 ($p<0.01$**). Rolling ROM and interaction percentage data for goals with toys on midline and sidewalls correlate with factors of 0.45 ($p<0.05$*) and 0.57 ($p<0.01$**), respectively. Grasping percentage and rolling ROM data for the goal with toys on sidewalls correlate with a Pearson coefficient R of 0.42 ($p<0.05$*). Correlation of FA orientation intensity and interaction percentage data was

additionally tested, resulting in R values of 0.61 ($p<0.01$**) and 0.60 ($p<0.01$**) for goals with toys on midline and arch, respectively. R values were determined also for the pair of FA orientation intensity and grasping percentage data and are 0.49 ($p<0.01$**) for midline and 0.60 ($p<0.01$**) for toys on the arch. Correlation of rolling ROM and max grasp pressure data is 0.67 ($p<0.01$**) for the sidewalls.

COP Movement in Sitting Position

Posture stability evaluation values for the sitting position trials are shown in Fig. 7. Data are given for

scenarios with and without toys along with statistical data comparison values. Statistically significant differences were identified for inter-goal comparison for the middle and higher groups, as well as consistently for the inter-group comparison. Correlation was estimated by taking into account the jointly included technically assessed stability parameter data of goals with and without toys and AIMS values. Coefficient R values are for the root-mean-square displacement parameter 0.58 ($p<0.01$**) and for the circle area parameter 0.56 ($p<0.01$**).

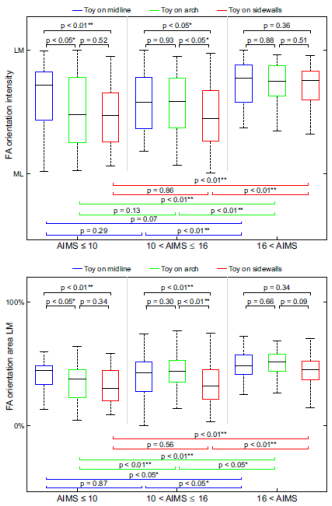


FIGURE 5. Forearm (FA) orientation intensity (upper subplot) and lateral-medial orientation area (lower subplot) values for goals with toys on midline (blue), arch (green), and sidewalls (red). The *p* values located above the coloured and black horizontal lines describe the inter-group and inter-goal data similarity comparison.

3.3.3 Discussion

The present work aimed to demonstrate that CareToy replies to the main requirements in the field of EI, allowing rehabilitation staff to remotely assess and stimulate preterm infants in the first year of life. The high and significant correlation values between the quantitative data of CareToy system and the AIMS scores, and the possibility of detecting different values between the three groups with different motor abilities (AIMS £ 10; 10<AIMS £ 16; 16<AIMS) demonstrate reliability of CareToy measurement data for discrimination of different infant's motor abilities. Moreover, the quantitative data determine specific assessments of different body segments (e.g. forearm orientation, body's rolling range of motion, grasp strength of the hands, etc.) and of different activities (e.g. toy interaction, postural stability) that cannot be detected by clinicians using only clinical scales. Regarding the motor abilities in supine position, CareToy data have shown that the group with higher abilities is significantly different from the other two groups. Comparison of these two groups has shown similar results, when reaching and grasping activities are required (toy on midline, toy on arch, and toy on sidewalls). As can be roughly detected by the items of AIMS, only infants with higher scores in supine position are able to roll for reaching and grasping activities, but in this case we quantitatively measured the differences in the body rolling ROM.

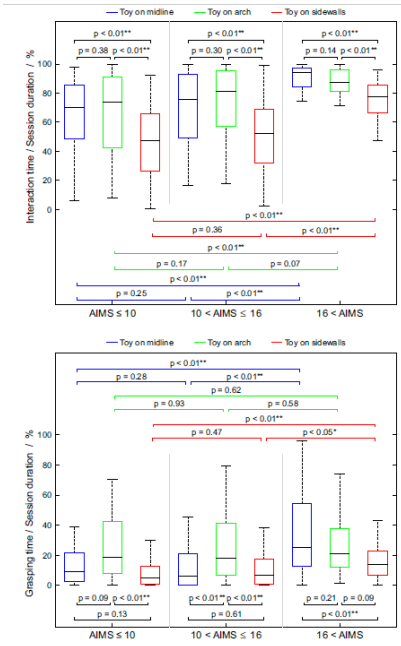


FIGURE 6. Toy interaction (upper subplot) and grasping (lower subplot) percentage values for goals with toys on midline (blue), arch (green), and sidewalls (red). The *p* values above the coloured horizontal lines describe the inter-group differences, while the ones above the black horizontal lines describe the inter-goal differences.

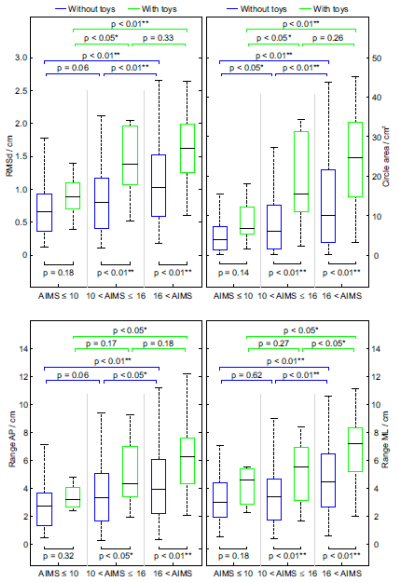


FIGURE 7. Sitting posture stability assessment results for goals with (green) and without toys (blue) in sitting position. Root-mean-square distance, circle area (both upper subplot), anterior-posterior and medial-lateral range (both lower subplot) of centre-of-pressure (COP) movement are given for infant groups, according to their AIMS values. The *p* values near the coloured horizontal lines describe the inter-group differences, while the ones next to the black horizontal lines describe the inter-goal differences.

In relation to the stimulation purpose of the CareToy system, results have indicated that all the infants, regardless of the AIMS-related group, orient their forearms significantly different, if they have the toy on midline or not (see Fig. 4). In fact, the presence of the toy on midline induces reaching behavior with the forearms towards the medial position, while whenever the toy is not present the forearms are oriented laterally. Reaching and grasping toys on midline (Toy on midline) is a fundamental goal to be promoted in the group of infants with lower and medium competencies, while it is already consolidated in infants with higher competencies. Rolling stimulation is intended for the lower group as head rolling, related also to the gaze movements, while it is for the other two groups more related to head and body movement. These two different clinical aims are confirmed by CareToy data, when comparing ROM values for the both goals and reviewing statistical analysis values. These reveal no significant differences in the group with lower abilities (only head rotation in sense of gaze

following was expected), and statistically significant differences for the other two groups (see Fig. 3). Comparing the other two assessed goals (Toy on arch and Toy on sidewalls), a difference in variability of the rolling ROM was expected and detected mainly in the first two groups of infants, because the toys, placed on the lateral walls, induced a higher displacement of the body. For the higher group, which is able in rolling, no differences were discovered, which was expected (see Fig. 3). Moreover, this was additionally confirmed by the forearm orientation intensity data that were more lateral in the lower group: this group does not roll a lot in view of reaching laterally positioned toys (arch and sidewalls), thus covering larger lateral area with the forearms (see Fig. 5). These goal-directed activities are mainly of reaching than grasping as confirmed by the data of interaction and grasping time that were significantly higher (mainly for reaching, detected by the interaction time) for toy on midline and on the arch (Fig. 6). Regarding the middle group an interesting finding, in accordance with the clinical aims, is that there were no significant differences in all parameters between the toy on midline and toy on the arch. These infants have acquired a good ability of reaching and grasping toys on midline, as well as crossing the midline for lateral grasping, showing same competencies in both goals (see Fig. 5). Another important finding are also the differences for the stimulation of reaching and grasping for small and high ranges that have in the last condition induced a significantly higher laterality of the forearm orientation intensity and area, which was however not followed by an increase of interaction and grasping time (see Fig. 6). The higher group showed a different strategy. In relation to their rolling ability, their forearm orientation remained medial in all the different aims of grasping, despite their ability of interaction and grasping of toys being lower in the extreme positions (toys on sidewalls) (see Fig. 6). The resulting statistically significant differences for the sitting position assessment, when comparing tasks that require only head and trunk control (without toys) and tasks that require reaching and grasping behavior (with toys) (see Fig. 7) are also very interesting. Middle and higher groups showed significant differences in all the computed parameters for

higher trunk movements, when toys were present, stimulating infants in moving their trunk and upper limbs towards the toys. The lack of differences in the group with lower competencies could be related to their very initial trunk control abilities, in which sense they perform very few attempts on moving their body towards the toys. This hypothesis is confirmed by the differences in the variability of trunk movements among the group with lower competencies and the other two groups, when the toys are present.

Study Limitations

The presented work had some specific study limitations. Training duration was absolutely limited to under 2 hours of training per day to avoid overburdening the infants' parents and loss of interest in the CareToy environment. Longer training duration would naturally result in more acquired sensor data that could perhaps enable more complex and reliable statistical analyses. On the other hand, the sole complexity of acquired data was the main reason for the need of such robust, complex, advanced data processing algorithms. Involvement of families in the rehabilitation process had several advantages, but also minor shortcomings. As the large part of rehabilitation process can take place at infants' home environment, parents can be present throughout training trials, and can by interaction additionally stimulate infants' activity and interest for training. This can however affect data processing, as parents occasionally stimulate infants by moving and shaking the CareToy sensorized toys, which can result in outliers and result inaccuracy. In the context of the presented study, this was however avoided by reviewing videos of all training trials and omitting the affected ones.

3.4 A randomized clinical trial in preterm infants on the effects of a home-based early intervention with the 'CareToy system'

CareToy system is an innovative tele-rehabilitative tool, useful in providing intensive, individualized, home-based, family-centred Early Intervention in infants. Our aim was to evaluate, through a Randomized Clinical Trial (RCT) study, the effects of CareToy intervention on early motor and visual development in preterm infants. 41 preterm infants (range age: 3.0±5.9 months of corrected age) were enrolled and randomized into two groups, CareToy and Standard Care. 19 infants randomized in CareToy group performed a 4-week CareToy program, while 22 allocated to control group completed 4 weeks of Standard Care. Infant Motor Profile (IMP) was primary outcome measure, Alberta Infant Motor Scale (AIMS) and Teller Acuity Cards were secondary ones. Assessments were carried out at baseline (T0) and at the end of CareToy training or Standard Care period (T1). T1 was the primary endpoint. After RCT phase, 17 infants from control group carried out a 4-week CareToy program, while 18 infants from the CareToy group continued with Standard Care. At the end of this phase, infants were re-assessed at T2. In RCT phase, delta IMP total score and variation and performance sub-domains were significantly higher ($P<0.050$) in CareToy group if compared to Standard Care group. Similar results were found for Teller Acuity Cards, while no differences between groups were found for AIMS. No differences were found in any outcome measure results (T2-T0), between infants who started CareToy training before or after one month of standard care. This RCT study confirms the results of a previous pilot study, indicating that CareToy system can provide effective home-based early intervention.

3.4.1 Materials and methods

Study design

This was a randomized, multicenter, evaluator-blinded, RCT, designed according to CONSORT Statement. At first, it was orginal designed as cross-over trial but after chosing T1 as the primary endpoint (see later) the trial was modified to a parallel RCT followed by an open phase. Eligible infants were identified by the two clinical centres involved in the study: IRCCS Fondazione Stella Maris with NICU of Pisa University Hospital in Italy and Helene Elsass Center with the University of Copenhagen in Denmark. Detailed information about the CareToy Project were provided to attending families. If parental written consent was obtained, infants were enrolled. Perinatal and clinical data were collected upon enrolment, and infants were evaluated during their first months through standard neurological examinations. At 3 months of corrected age, infant gross-motor abilities were quantified by the Ages & Stage Questionnaire Third Edition (ASQ-3) gross motor area, filled out by parents, in order to define the most appropriate starting point for each infant. After baseline evaluation (T0), infants were randomly assigned to either CareToy group or Standard Care group. Twins were allocated to the same group. All infants were re-assessed at T1 (4 weeks post T0), on primary and secondary outcomes. After this RCT phase, according to study design, infants who started out with Standard Care at T1 were given the opportunity to carry out a 4-week CareToy program, while infants that had already performed CareToy training switched over to Standard Care. 4 weeks after this phase all infants were re-assessed (T2). A further assessment is planned at 18 months of corrected age (T3), but the results of this timepoint are not included in this paper. All assessments were recorded on video and scored offline by expert therapists, blinded to group allocation.

The study was conducted according to Good Clinical Practice and Declaration of Helsinki principles and supervised by an Ethics Advisory Board. Approval

was obtained by the Ethics Committee of the Hovedstaden Region (Denmark) in June 2012, Ethics Committee of Pisa University Hospital (Italy) in June 2013, Tuscan Region Pediatric Ethics Committee (Italy) in February 2014 and IRCCS Fondazione Stella Maris Review Board in May 2014. The first infant was enrolled in the pilot study in July 2013, but stopped afterwards due to CareToy technical adjustments and restarted in November 2013 when the protocol was registered. The authors confirm that the trial has been registered at ClinicalTrials.gov (NCT01990183, https://clinicaltrials.gov/ct2/show/NCT01990183).

Study populations

Eligible infants were preterm infants with the following characteristics:

- born between 28 + 0 and 32 + 6 (weeks + days) of gestational age;
- aged 3±9 months of corrected age who had achieved a predefined cut-off score in gross motor ability derived from ASQ-3.

Exclusion Criteria were:

- birth weight below the 10th percentile;
- brain damage i.e. intra-ventricular haemorrhage < grade 1, any degree of periventricular leukomalacia, or brain malformation;
- any form of seizure;
- severe sensory deficits (blindness, deafness);
- other severe non-neurological malformations;
- participation in other experimental rehabilitation studies.

Intervention conditions

CareToy intervention. CareToy intervention (detailed in Sgandurra et al) is an intensive, highly customized, home-based, family-centred training program, provided through remote management of a CareToy system delivered at home. It consists of specific goal directed activities, called scenarios, remotely planned by the clinical/rehabilitative staff according to specific infant needs and capabilities. Training is multiaxial with a high degree of variability and

complexity. Each scenario lasts from 2 to 10 minutes and promotes different aspects of motor and visual development, such as head rotation and gaze movement, grasping and eye-hand coordination. Based on postural control and rehabilitative needs, activities integrated in CareToy system can be variably planned in supine, prone or sitting position. Training is programmed daily for 30±45 minutes for 4 weeks including weekends (a total of 28 days). According to general guidelines previously developed during pilot study (16), CareToy training at home is organized into two phases. The first one, lasting a week, is devoted to infant and parent habituation to system and identification of main rehabilitation goals. The second phase, lasting three weeks, consists of continuous planning and customization of training in relation to daily activities and progress. At the end of each day, the CareToy System automatically sends a report of training to rehabilitative staff, so that it can be monitored and directed to promote progressively more complex abilities when the previous ones have been achieved.

Standard care. Standard Care consists of a bimonthly follow- up check, during which current care advice on the early management of preterm infants and booklets dedicated to home-care of preterm infants are distributed, according to the standard recommandations of the two involved countries (Italy and Denmark). In rare cases, sporadic sessions with a physical therapist for special assistance were arranged. In such cases, the number and type of performed activities were recorded.

Measures

Due to the characteristics of CareToy intervention which mainly address goal-directed motor activities, a motor scale as primary outcome measure was chosen. In particular, the Infant Motor Profile (IMP), previously indicated as secondary outcome, was shifted to the primary one. The study protocol, as registered at Clinicaltrials.gov, was modified accordingly and published in the description of study protocol and in the results of pilot study. IMP is a recent test aimed at assessing motor behaviour of preterm and term infants aged 3±18

months. It consists of 80 items scored off-line on the basis of standardized video recordings. Total IMP score constitutes the mean of five subdomains (variation, adaptability, symmetry, fluency and performance). Adaptability is only scored for infants older than 6 months, so it was not included in this study. This evaluation was carried out at T0, T1 and T2. Secondary outcome measures consisted of one for motor and one for visual assessments. The first one was the Alberta Infant Motor Scale (AIMS). It is a standardized scale, used in infants from term until 18 months of age. It assesses infant motor abilities and quality of posture and movement in four positions: prone, supine, sitting and standing. It is possible to determine a total overall score and subscores for each assessed position. AIMS evaluation was performed at T0, T1 and T2. At the same time points, Teller Acuity Cards were used to evaluate visual acuity. It is based on assessment of infant attention to a series of cards showing stripes of different widths. This tool allows for a rapid assessment of visual acuity (grating) in infants and young children and other populations where verbal response to recognition of visual acuity charts (letters) is difficult or impossible. It evaluates development of visual acuity and has been used in several studies for diagnostic purposes and to assess results of early intervention.

Statistical analyses

Sample size was calculated on the basis of previous results on IMP and on design of a parallel RCT study. 38 infants (42 in total, including drop-outs) were needed to detect a clinically relevant change of 2.4 points with a power of 80% at a significance level of 0.05. Clinical data were analyzed by means of Statistical Package for Social Sciences (SPSS, version 20.0). Normality of distribution was verified by Shapiro-Wilk's test. To test a-priori baseline differences between the two groups for characteristics (Gestational age, Birth weight, Gender, Corrected Age) and baseline measures (IMP total and subdomains, AIMS total and subdomains and Teller) the t-test for unrelated samples and non-parametric Mann-Whithney test were used for normal and

non-normal distributed data, respectively. Programmed and executed days and hours of CareToy training and their rate were calculated in order to assess compliance with CareToy intervention. In particular, rates were calculated by computing percentages of ratio between executed and programmed days and hours, respectively. Changes in primary and secondary outcome measures were calculated between baseline (T0) and post-intervention period (T1). In order to verify effect of CareToy training versus Standard Care on primary endpoint (T1-T0), the t-test for unrelated samples and the non-parametric Mann-Whithney test were used for normal and non-normal distributed scores, respectively. Effect size was computed using Cohen's d for IMP total, AIMS total and Teller. Commonly used criteria specify that a value below 0.2 is regarded as no effect, a value of 0.2±0.5 as a small effect, a value of 0.5±0.8 as a medium-sized effect and a value above 0.8 as a large effect. Moreover, a Pearson correlation analysis was carried out between AIMS total delta scores (T1-T0) and IMP total and IMP performance delta scores (T1-T0), respectively, in order to detect if there were any correlations between the two motor outcome measures, as detected in the validation studies of IMP. A further analysis was planned to evaluate whether CareToy training carried out one month after of standard care (T1) may have different effects than CareToy performed at the beginning (T0). In particular, we compared the delta of T2-T0 between the two groups. Finally, correlations between delta changes after CareToy training in primary and secondary outcome measures that showed statistical significant changes and hours of CareToy training performed were determined by linear regression analysis (Pearson correlation coefficient). Non-adjusted significance level for all analyses was set at $p<0.05$.

3.4.2 Results

Participants

249 infants were assessed for eligibility. 208 were excluded from study because they did not

meet inclusion criteria (160) or families declined to participate (48) (Fig 1). 41 infants (mean age 3.9 ± 0.8 months, range: 3.0±5.9) were recruited between May 2014 and April 2015 and 19 infants (mean age 3.6 ± 0.4 months, range: 3.1±4.9) were randomly allocated to CareToy and 22 (mean age 4.1 ± 1.0 months, range: 3.0±5.9) to Standard Care. Baseline characteristics of participants are reported in Table 1. No statistically significant differences were found between the two groups at baseline. All infants underwent 4 weeks of intervention according to their allocation. Infants allocated to the CareToy group received a mean of 25.6 ± 3.1 days and 10.7 ± 2.1 hours of training with a ratio to programmed days and hours of 95.1% and 72.1%, respectively. None of the infants of CareToy or Standard Care group received any special sessions with a physical therapist. All infants were reassessed the week after the end of intervention period (T1).

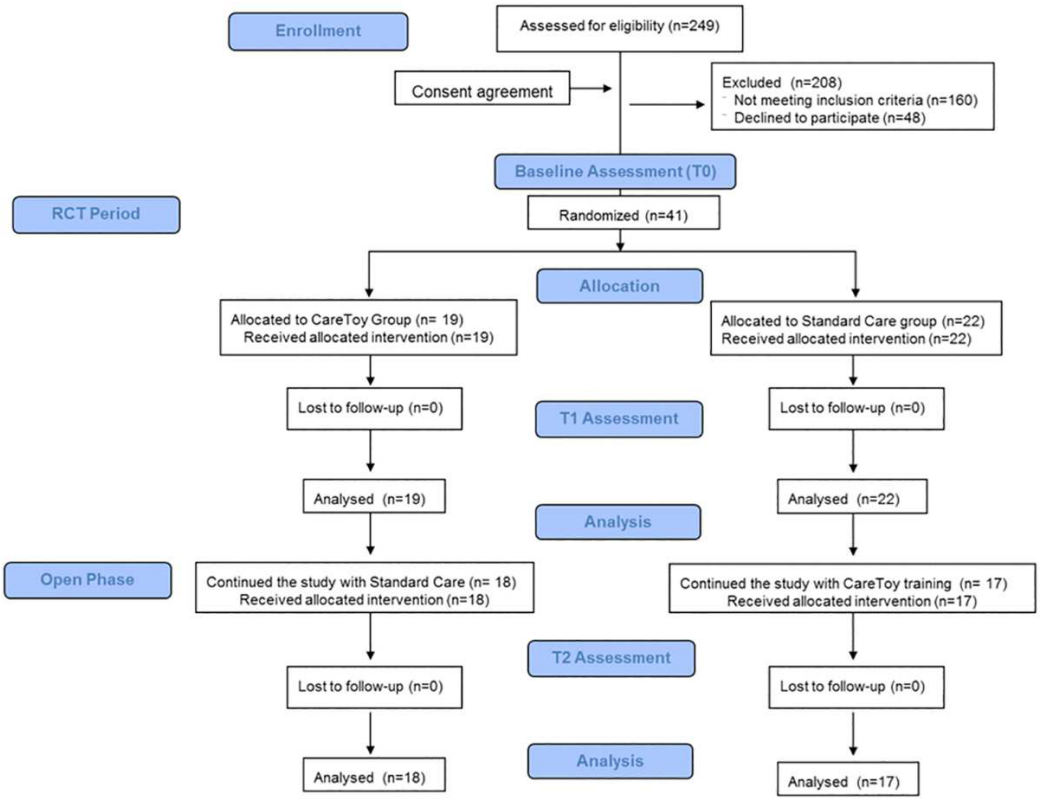


Fig 1. Flow-chart of study according to CONSORT diagram.

After this RCT period, 17/22 infants previously allocated to the Standard Care group carried out CareToy training for 4 weeks (mean of 21.9 ± 4.2 days and 9.9 ± 3.6 hours with a ratio to the programmed days and hours of 95.6% and 79.6%, respectively) and were reassessed at T2. At the same time point (T2), 18/19 infants previously allocated to the CareToy group were re-evaluated after the Standard Care period.

Considering the whole group of 26 infants who carried out the CareToy intervention in the RCT period (19) and in the open phase (17) the infants received a mean of 25.1 ± 4.0 days and 13.7 ± 3.8 hours of training with a ratio to programmed days and hours of 95.3% and 75.3%, respectively.

Primary outcome measure (IMP)

After 4-week intervention period, delta IMP total score (change from T0 to T1) was significantly higher in CareToy group compared to Standard Care group

(Table 2). Moreover, there were also significant changes for Performance and Variation subdomains. The Cohen's d value was 0.69. There were no differences in changes between the two groups from T2 to T0 (Table 3).

Secondary outcome measures

AIMS. After 4-week intervention period, change was higher (but not statistically significant) in CareToy group compared to Standard Care for AIMS total score and also for supine and prone subdomains (Table 2). Cohen's d value of AIMS total was 0.33. Pearson correlation analysis showed a significant correlation between changes in AIMS total delta scores and IMP total (0.572, $p<0.0001$) and IMP performance (0.597, $p<0.0001$) delta scores of T1-T0.

There were no differences in changes between the two groups from T2 to T0 (Table 3).

Teller acuity cards. After 4-week intervention period, there was a significantly greater improvement in visual acuity in CareToy group compared to Standard Care group (Table 2) with a Cohen's d value of 0.68.

Table 1. 10 Main characteristics of the CareToy and Standard Care groups at baseline (T0).

Characteristic		CareToy group (n = 19)	Standard Care group (n = 22)	*p value*^
Gestational age (weeks), mean (SD)		30.7 (1.4)	30.82(1.1)	0.772 *
Birth weight (grams), mean (SD)		1368.3 (330.3)	1459.5 (275.6)	0.308§
Gender M/F		8/11	11/11	
Corrected Age (months), mean (SD)		3.6 (0.4)	4.1 (1.0)	0.704§
ASQ-3 Gross Motor Area				
	Form 2 (median)	n = 3/19 (40)	n = 2/22 (52.5)	-
	Form 4 (median)	n = 16/19 (50)	n = 16/22 (50)	-
	Form 6 (median)	-	n = 4/22 (42.5)	-
Twins		8/19	10/22	
Infant Motor Profile, mean (SD)				
	Total	69.9 (3.5)	69.8 (3.4)	0.924*
	Performance	45.9 (6.6)	48.1 (7.4)	0.239§
	Variation	64.9 (5.2)	63.6 (3.0)	0.820§
	Fluency	75.0 (0.0)	75.0 (0.0)	1.000§
	Symmetry	93.8 (8.1)	92.5 (7.8)	0.387§
AIMS, mean (SD)				
	Total	10.8 (3.6)	11.7 (3.3)	0.316§
	Prone	3.5 (1.4)	3.9 (1.5)	0.583§
	Supine	4.4 (1.0)	5.0 (1.3)	0.121§
	Sitting	1.7 (1.6)	1.5 (1.0)	0.731§
	Standing	1.2 (0.5)	1.3 (0.4)	0.445§
Teller Acuity Card (cy/degree), mean (SD)		2.6 (1.0)	3.2 (1.2)	0.158§

Abbreviations: n: number; SD: Standard Deviation

^ Significant non-adjusted level <0.050

*Independent t test

§ Mann-Whitney

Regression analysis

Regression analysis carried out immediately after training showed a significant relationship between delta changes in Teller scores and amount of performed CareToy training ($r = 0.411$, $p = 0.013$) while it was not significant for delta changes in IMP scores ($r = 0.183$, $p = 0.285$) (Fig 2).

3.4.3 Discussion

Motor and visual development are frequently delayed in first months of life in preterm infants. EI programs are aimed at minimizing these developmental delays. In this framework, the main goals of CareToy training are to promote postural control, reaching, grasping, visual attention and orientation in highly variable and tailored activities set at various levels of complexity. Results of RCT period (T1 versus T0) indicate that the 4-week EI program provided through CareToy system can significantly improve motor and visual development in preterm infants. In the motor domain of the primary outcome measure (IMP), we found medium sized, statistically significant effects on IMP total scores compared to Standard Care. The effect appeared to be driven by improvements in variation and performance subdomains.

Table 2. Mean score increase from before 4-week intervention to after (T1-T0).

		CareToy group (n = 19)	Standard Care group (n = 22)	*p value*^
		Delta mean (SD)	Delta mean (SD)	
Primary outcome measure				
IMP				
	Total	6.0 (2.2)	4.3 (2.7)	0.039*
	Performance	10.6 (4.3)	7.3 (5.0)	0.028*
	Variation	8.3 (6.6)	6.9 (5.2)	0.044*
	Fluency	1.1 (3.4)	0.00 (0.00)	0.123§
	Symmetry	3.7 (6.6)	4.4 (6.0)	0.440§
Secondary outcome measures				
AIMS				
	Total	4.8 (3.1)	3.9 (2.7)	0.308*
	Prone	2.2 (2.1)	1.8 (1.5)	0.749§
	Supine	1.7 (0.9)	1.0 (1.0)	0.077*
	Sitting	0.5 (0.8)	0.6 (1.4)	0.669§
	Standing	0.4 (0.5)	0.4 (0.5)	0.939§
Teller Acuity Card (cy/degree)		2.0 (1.2)	1.2 (1.4)	0.035*

Abbreviations: n: number; SD: Standard Deviation

^ Significant level <0.050

*Independent t test

§ Mann-Whitney test

https://doi.org/10.1371/journal.pone.0173521.t002

Table 3. Mean increase of T2-T0 in the two groups.

		CareToy + plus SC (n = 18)	Standard Care + CareToy (n = 17)	*p value*^
		Delta mean (SD)	Delta mean (SD)	
Primary outcome measure				
Infant Motor Profile				*
	Total	9.5 (3.1)	9.2 (3.6)	0.091*
	Performance	15.9 (5.0)	15.4 (4.9)	0.077*
	Variation	14.1 (6.6)	14.3 (5.5)	0.092*
	Fluency	1.1 (3.4)	0.5 (2.0)	0.563§
	Symmetry	3.9 (7.6)	7.1 (9.5)	0.115§
Secondary outcome measures				
AIMS				
	Total	9.4 (4.0)	9.6 (3.2)	0.870*
	Prone	4.1 (2.7)	4.5 (2.0)	0.615*
	Supine	2.3 (1.1)	2.6 (0.9)	0.380*
	Sitting	2.2 (1.8)	1.8 (1.6)	0.586§
	Standing	0.8 (1.0)	0.5 (0.5)	0.273§
Teller Acuity Card (cy/degree)		3.2 (1.3)	3.5 (1.4)	0.361*

Abbreviations: n: number; SD: Standard Deviation

^ Significant level <0.050

*Independent t test

§ Mann-Whitney test

Variation is brought about by the explorative activity of nervous system and is used for shaping it, reflecting the interaction between individual motor repertoire and environment. A high variation indicates an abundance in cerebral connectivity in typical development, while a more limited variation can be accounted for by structural anomalies in which disturbances of cortical connectivity may play a prominent role. Enlargement of limited movement repertoire is an aim of EI, even if animal data indicate that this is difficult to achieve. To our knowledge, EI through CareToy is the first program that has

demonstrated significant effects, when formally compared to Standard Care, in enhancing variability of motor behavior. In addition, CareToy EI is able to promote an earlier achievement of developmental motor milestones evaluated by the IMP performance subdomain. Fluency (i.e. ability to fine-tune motor output) and symmetry (i.e. presence or absence of stereotyped asymmetries) subdomains were not significantly different in CareToy group if compared to Standard Care one. However, the CareToy seems to be able to enhance fluency, whereas no changes were detected in Standard Care group. Although CareToy system has the ability to stimulate goal-directed activities for asymmetric conditions, the lack of effect on symmetry scores is likely explained by high symmetry scores at baseline, indicating that the sample of infants of this study did not need this kind of intervention.

Regarding the secondary motor outcome measure (AIMS), we found only a slight effect in the total score, with higher changes in supine and prone scores in the CareToy group compared to the Standard Care one. These two positions, supine and prone, were the most trained ones in relation to the age window and developmental needs of enrolled infants. Lack of significant effects of Care Toy training on AIMS measures, unlike IMP, could be related to the structure of AIMS items and relative scoring. Each item within the ªmotor windowº of AIMS is scored as “observed” or “not observed”, so minor changes cannot be specified. In addition, AIMS only assesses gross motor development while IMP evaluates both gross and fine motor development. Another aspect could be related to the different psychometric properties of the two scales and limitations for detecting short term effects by AIMS, as already pointed out in literature. In any case, a significant correlation was found between changes in AIMS total scores and IMP total and performance scores.

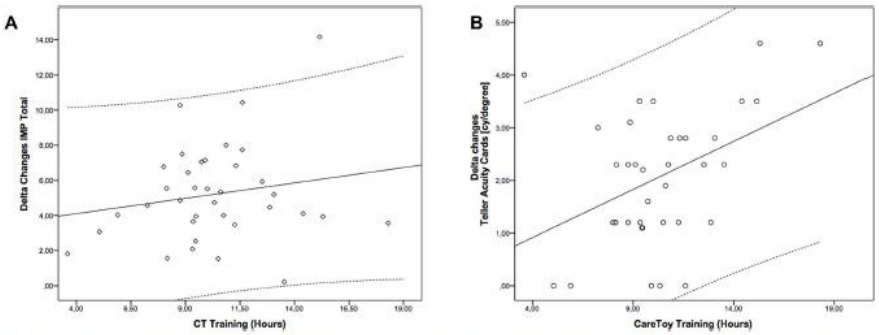


Fig 2. Regression analysis. Legend. Scatter plot of delta changes after CareToy training in Infant Motor Profile (A) and in Teller Acuity Cards versus number of hours of performed CareToy training. Solide lines represent the linear fit of data. Dotted lines represent Confidence Interval at 95%.

A further interesting result is that EI provided through the CareToy system is able to significantly promote visual acuity as assessed by Teller Acuity cards, similar to other EI programs in preterm infants, such as infant massage therapy. This finding is probably related to the high sensitivity of visual development in the first months, characterized by rapid maturation of visual cortex and visual acuity. Promotion of visual development is particularly crucial in preterm infants because a significant number show delayed maturation of visual acuity, compared to their term-born peers, with further deficits in domains of visual perception and visual-motor integration. Thus, early promotion of visual acuity can have a positive effect also on these more complex visual abilities. Moreover, considering the whole group of infants that performed CareToy training, we found a positive correlation between visual acuity changes and number of hours of training, while no correlation was found for IMP changes. This finding could be related to differences in dosage response between visual and motor system, as well documented in literature, where a relationship between dosage and effects has been found for the treatment of amblyopia and no relationship between dosage and effects for motor treatment. Moreover, our finding could be related to the type of measurement and type of training. The goal-directed motor activities promoted in CareToy training are largely dependent on the character and preferences of each infant (e.g. from early reaching on midline to grasping while rolling) and IMP scale is able to capture several aspects of motor development. This variability can justify why we found improvement only in the whole group for

IMP scores. On the contrary, Teller Acuity cards evaluate a specific aspect of visual development, i.e. visual acuity, that was specifically trained for in all infants using visual stimuli (e.g. videos on screen, lights on arch, on toys and on feedback walls). Moreover, no differences were found for any of the outcome measures, when comparing the results of CareToy training carried out before or after one month of standard care. These results justify the ethical decision of offering CareToy training to those infants who were randomly allocated in the Standard Sare group. All these data confirm the hypothesis that Care-Toy training when performed for 4 weeks in the very first months is able to have at least a short term impact on motor and visual development. The choice to have an open phase after the RCT period, creditable from an ethical point of view, hampers the evaluation and extrapolation of longer term effects of CareToy training. A strong relation has been reported between visual and motor functions in the first months of life and infant development in the other domains and long-term neurodevelopmental outcomes of newborn infants at risk. Moreover, results of RCTs and systematic reviews, indicate that EI has positive effects on cognitive and motor development, and hopefully could reduce prevalence of neurodevelopmental disabilities and improve quality of life.

3.5 Sensorized toys for measuring the effect of home-based therapy on manipulation capabilities of infants

Grasping is one of the main motor abilities that infants develop in the first months of life and it is a fundamental landmark in infants' development, allowing them to explore and learn about the surroundings. An efficient object grasping requires a number of skills (differentiation of individual fingers' movement, grading of grip force, control over the release), and recent studies confirmed that the measurement of grasping actions can be used to detect early signs of neurodevelopmental disorders.

Current studies on infants' motor development including grasping are mainly based on functional scales or on the direct observation of the infant while playing. Quantitative variables are rarely measured in clinical practice.

On this background, in this work we present and discuss the validation of a kit of sensorized toys as a useful tool for measuring preterm infants' manipulation capabilities and the effect of therapy given as home-based daily training. The kit of sensorized toys is part of a more complex system, a smart baby gym, developed in the framework of the CareToy Project (http://www.caretoy.eu). The measurements were acquired during a randomized controlled trial (RCT) of the baby gym, performed with 26 infants receiving daily home training for 4 weeks, with purposive protocols tailored on the infants' needs and development

3.5.1 Materials and methods

Experimental set-up

The complete CareToy baby gym for EI is described in detail in previous sub-chapters. The rehabilitative staff involved in the study used the modular organization of the baby gym to tailor the daily training sessions according to individual needs, thus stimulating the infant to perform specific goal-directed

activities (e.g. postural control, reaching, grasping, visual attention and orientation). In this paper, we focused on the training sessions designed to stimulate the grasping abilities on midline, involving the kit of sensorized toys (34)-(36).

The three toys, purposively developed for the baby gym, were inspired by commercial toys for infants (Figure 1, left). Their designs were studied on the basis of affordance (37) to encourage different manipulation approaches and to be compliant with the infants' hand dimensions (38). Custom molds were developed for manufacturing the graspable parts made of elastomer and realized through injection molding.

Toy hardware includes pressure sensors, IMU board with microcontroller, force sensors, and multicolor LEDs for light feedback. Sensors were chosen considering the range of force exerted by the infants in the first year of life (0-35 *kPa* and 0-20 *N*). The microcontroller interfaces sensors and lights embedded in the toys. Signals are acquired via USB and saved on the laptop for post-processing.

Figure 1. The CareToy platform (left). The infant in supine position is grasping one of the sensorized toys that is hung on the arch during a training session (right).

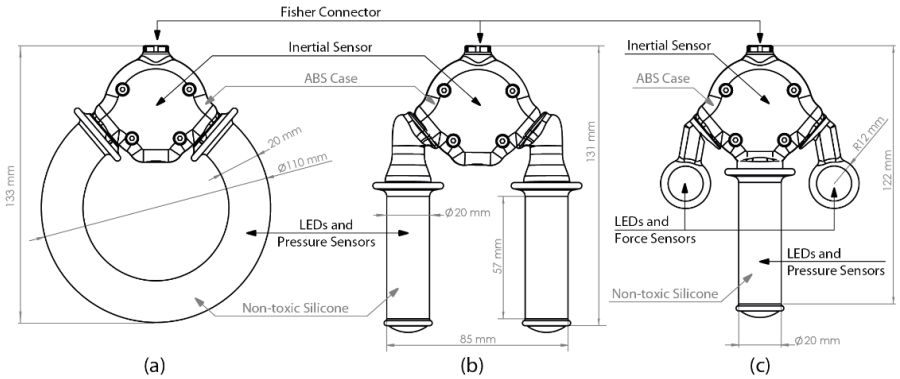


Figure 2. CareToy kit of toys: a) the large ring, used to encourage a (unimanual or bimanual) palmar grasp; b) the U-toy, used to induce a bimanual palmar grasp on each side of the toy; c) the Mouse-Toy, that promotes a palmar grasp on the cylinder and a pincer grasp on the two small "ears". Each toy is composed of a rigid part in ABS with embedded electronic board (this is comprised of the iNemo, ST Microelectronics, IT) and a soft sensitive part (nontoxic silicone RTV 4428 Silbione®, Bluestar Silicones, USA). The soft silicone chambers are connected to a pressure sensor (LPS331AP, ST Microelectronics, IT), force sensors (FSR®, Interlink Electronics, USA). The connection with the laptop is guaranteed by a push up aluminum connector (model DBPC 102 A053-13, Fischer Connectors, USA). Cables were chosen considering the safety requirements in terms of material and cable length (200 mm to avoid strangulation).

Three basic shapes have been chosen for the sensorized toys, on the basis of the type of grasp (palmar or digital) and hand action (unimanual or bimanual) to detect and measure:

1. large ring (Fig. 2(a)): a toroidal pressure chamber equipped with pressure sensors and LEDs inside;

2. U-Toy (Fig. 2(b)): a horseshoe composed of two cylindrical pressure chambers with a pressure sensor and lights in each side;

3. Mouse-Toy (Fig. 2(c)): a cylinder equipped with a pressure sensor and LEDs inside the main body and two small "ears" embedding a force sensor and a LED.

Clinical sample and training sessions

The study involved preterm infants meeting the following inclusion criteria:

i) gestational age ≥28 weeks and ≤32 weeks;

ii) corrected age (CA, i.e. the age of a preterm infant that is starting from the expected date of delivery and that is calculated by subtracting 40 weeks of gestation from the actual date of birth) at baseline between 3 and 9 months;
iii) adequate in terms of milestones in gross-motor abilities.

Infants with brain damage i.e. brain malformation, intra-ventricular haemorrhage (IVH) >1, any degree of periventricular leukomalacia (PVL), known epilepsy, severe sensory deficits (blindness, deafness), or other severe non neurological malformations were excluded.

Details regarding the clinical protocol are deeply described in (33). The primary outcome measure was the Infant Motor Profile (IMP): a video-based clinical scale composed of 80 items for the assessment of motor behaviour in infancy. In particular, several items are devoted to the measurement of reaching and grasping in different positions (i.e. prone, supine and sitting). Moreover, IMP has showed a promising predictive ability on differentiating typical from atypical development (39)-(40).

The study has been approved by the Ethics Committee of Pisa University Hospital (Italy) and Tuscan Region Pediatric Ethics Committee (Italy), and it has involved the clinical center IRCCS Fondazione Stella Maris, Department of Developmental Neuroscience, in Pisa (Italy), in collaboration with Neonatal Intensive Care Unit, Pisa University Hospital "Santa Chiara". The trial was registered at ClinicalTrials.gov (NCT01990183). All the parents of the enrolled infants provided their written informed consent to enter the study.

The current trial comprised 26 preterm infants with CA between 13.3 and 26.3 weeks and IMP score between 64 and 80 at the beginning of the training; each subject performed 4 weeks of daily training at home.

The mechatronic toys were employed in the training to promote grasping activities while gathering quantitative information. During each session, one of the three toys was hung on the midline of the arch while the infant was in supine position (Figure 1, right). The blinking of the LEDs was expected to encourage the infant in playing with the toy. Each training session lasted about three minutes, for an average overall training time of 53 min (±28 min); each

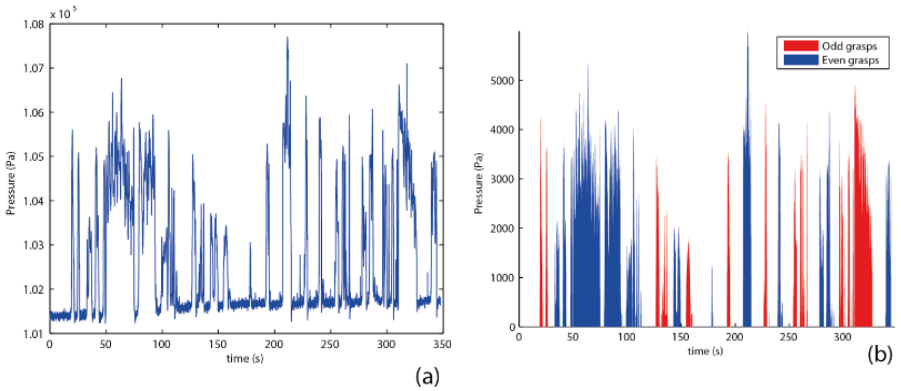
subject performed 4.43 training sessions per week on average. The total amount of the analyzed training sessions was 468.

Manipulation parameters analysis

To obtain a measure of the infants' manipulation capabilities and in particular palmar grasp activity, we analyzed signals from the pressure sensors (generated by palmar grasps) embedded in each toy. Inertial and force sensors data are not reported in the present work since they regard other specific aspects of manipulation capabilities (e.g. precision grasp and arms dynamics). Raw pressure signals have been post-processed by filtering and removal of artefacts (Figure 3). Such processed signals have been employed to calculate three significant parameters, independently from the shape of the toy used in each session:

1. Grasping time over training time: dimensionless, it is the ratio between the time the infant is grasping the toy and the total training time;

2. Mean grasp length (*s*): it is the mean interval between beginning and end of a single grasp. Grasping actions are assumed to start when the pressure value



exceeds a triggering threshold (set at 100 *Pa*) for more than 1 s. The grasp is considered to be ended when the pressure signal keeps below the threshold longer than 3 *s* (Figure 3(b));3. Mean grasp intensity (*Pa*): it is the average pressure recorded by the pressure sensor during the grasping activity

Figure 3. Example of pressure signal from a toy recorded during a training session: a) the raw signal; b) the post- processed signal after filtering and removal of the atmospheric pressure

baseline and other artifacts. In particular single grasping actions are identified and isolated when the pressure signal stands over a threshold of 100 Pa for more than 1 s. Consecutive grasps of a sequence are highlighted in the plot with two different colors (blue and red peaks for even and odd grasps respectively).

Data collected during the training sessions have been organized in an object-oriented database, defining a session class whose member functions allow execution of the required analysis over single objects as well as groups.

We computed then the three parameters on the first and last three training sessions of each infant. Data of the 26 infants were grouped according to their motor performances.

Infants were uniformly divided into three groups on the basis of their IMP score:

- IMP < 72, defined as "Low" group;
- 72 ≤IMP ≤ 74, defined as "Medium" group;
- IMP > 74, defined as "High" group.

With the aim of considering the "CareToy effect" (i.e. expected effect of training on infants' motor development) we evaluated two developmental trends:

a) a baseline trend, obtained considering data averaged on the first three sessions and reported in correspondence of the infants' IMP score at the beginning of the study (Low/Medium/High Start data). This information allows to extract a potential trend expected for an infant who did not receive the CareToy training;

b) CareToy trend, obtained by applying the same procedure as for a) to the last three training sessions of each subject. This information can be interpreted as the developmental trend of an infant provided with the EI treatment.

Statistics

Non-parametric analysis using the Wilcoxon and Mann-Whitney tests was accomplished to look for statistically significant differences in the previously

described parameters, comparing the groups of subjects described above. P-values less than 0.05 highlight statistically significant differences between two sets of data.

The search of statistically significant variations was done comparing:

a) the starting motor competences of the three groups (Low Start vs. Medium Start, Medium Start vs. High Start, and Low Start vs. High Start);

b) the CareToy effect over each group by contrasting the motor parameters at the beginning and at the end of training (Low/Medium/High Start vs. Low/Medium/High Stop respectively). Moreover, in order to distinguish between the actual training outcome and the natural development, a further comparison was achieved between the training end of a group and the training beginning of a group with superior IMP scores (comparisons done: Low Stop vs. Medium Start, Medium Stop vs. High Start, and Low Stop vs. High Start). Indeed, according to (41), we expect a clinically relevant change of 2.4 points in the IMP score due to the training.

3.5.2 Results

In this section, boxplots of data are reported for the three parameters. Results regarding the developmental aspects under investigation are presented in accordance to the above mentioned statistical analyses.

Grasping time over training time

Figure 4 shows the boxplot of the values of Grasping time over training time for the three infants' groups of different IMP at the beginning and at the end of the training. Median values are reported in Table 1.

Parameter	Low Group		Medium Group		High Group	
	Start	Stop	Start	Start	Start	Stop
Grasping time over training time	0.23	0.38	0.15	0.38	0.38	0.49
Mean grasp length (s)	11.75	13.60	10.21	20.22	12.46	20.75
Mean grasp Intensity (Pa)	1344.2	3057.0	1074.8	2260.0	1536.2	2837.6

Table 1. The median values of the distributions for the three infants groups of different IMP at the beginning and at the end of the training for the three selected parameters

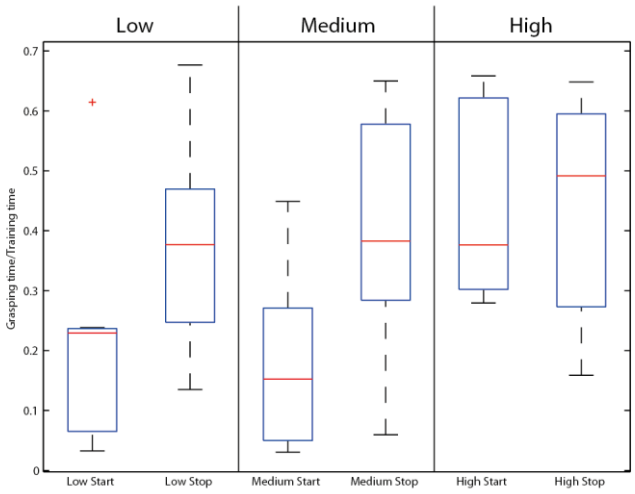


Figure 4. Boxplot of Grasping time over training time values for the three infants' groups of different IMP at the beginning of the training. Data were obtained considering their first and last three sessions (Low/Medium/High Start and Low/Medium/High Stop, respectively)

a) Statistically relevant differences are revealed from the comparisons of the distributions of Grasping time over training time, in the Low Start and High Start groups (p Low Start vs High Start=0.003) and Medium Start and High Start groups (p Medium Start vs High Start=0.003).

b) Start and Stop distributions show significant differences in case of Low and Medium groups (p Low Start vs Low Stop=0.03, p Medium Start vs Medium Stop=0.01). The Low Stop distribution is different from the Medium Start one (p Low Stop vs Medium Start=0.03), but not from the High Start one (p Low Stop vs High Stop median=0.37).

Mean grasp length

Figure 5 shows the Boxplot of Mean Grasp length values for the three infants' groups of different IMP scores at the beginning and at the end of the training. Median values are reported in Table 1.

Performed comparison detected no significant differences for the Mean grasp length analysis.

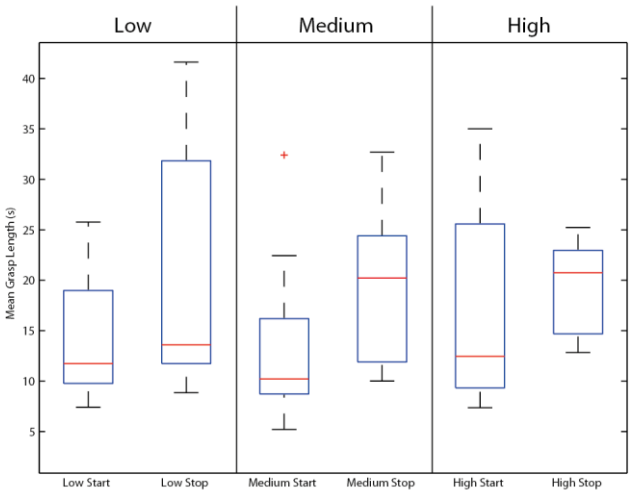


Figure 5. Boxplot of Mean grasp length values (*s*) for the three infants' groups of different IMP at the beginning of the training. Data were obtained considering their first and last three sessions (Low/Medium/High Start and Low/Medium/High Stop, respectively)

Mean grasp intensity

Figure 6 shows the Boxplot of Mean Grasp intensity values for the three infants' groups of different IMP scores at the beginning and at the end of the training. Median values are reported in Table 1.

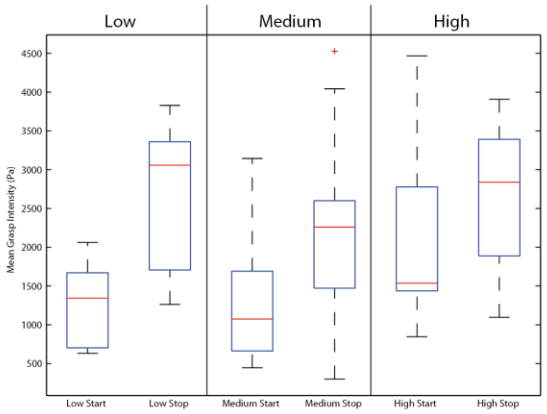


Figure 6. Boxplot of Mean grasp intensity (*Pa*) for the three infants' groups of different IMP at the beginning of the training. Data were obtained considering their first and last three sessions (Low/Medium/High Start and Low/Medium/High Stop, respectively)

a) The baseline developmental analysis did not show significant differences.
b) Comparisons between Start and Stop distributions were statistically significant for the Low group (p Low Start vs Low Stop=0.005) and the Medium group (p Medium Start vs Medium Stop=0.04). The Low Stop group differs from the Medium Start one (p Low Stop vs Medium Start=0.003), but it is comparable to the High Start group.

3.5.3 Discussion

The objective of the current work was to evaluate the measurements done by the kit of the CareToy sensorized toys during clinical trials of home-based daily training. The toys could monitor in a quantitative way the babies' grasping actions during the personalized training sessions at each infant's home, tailored by varying the kind of toys and the stimuli. The working range and the sensitivity of the toy sensors resulted to be adequate for the age of the sample and the selected parameters were proved to be good indicators to track differences and trends along the training sessions.
Regarding the clinical validation, the sensorized toys detects the CareToy positive effect on the ability of exploring toys in the presented sample.
Indeed, literature findings show that preterm infants spend less time in exploring objects (23) and this behavior is supported from the grasping time over training time parameter. This parameter confirms that High group infants spent more time in grasping the toys.
This is further highlighted if we look the not significant results of grasp length. In fact, the number of grasps in a training session increased along the CareToy training (i.e. if the time spent in playing with the toys increased and the length remained unchanged, the number of grasps in a training session should increase consequently).

Moreover, if we analyze the baseline trends for the grasping time over training time values, the Low and Medium groups have similar values, both lower than the last group. After the CareToy, the values measured in the Low group are much higher, even higher than the Medium Start and equal to the High group. Concerning the grasp intensity, the toy measurements clearly show the CareToy significant effect in increasing the grasp strength in the Low and Medium groups, especially considering that the developmental trend shows no differences among the three groups even if the High group has more spread data covering high values.

There is a methodological limitation of the current study that is important to highlight, relying in the data analysis: the toys were considered as a whole since the variability of the CareToy training (tailored on the infants' needs and on the subgoals of the grasping on the midline, e.g. unimanual or bimanual) did not allow to have enough data for the analysis of each single toy. However, the present work represents an innovative approach for measuring infants' grasping abilities in a quantitative way and assessing the training effects. This approach could have a huge impact in the clinical evaluation and remote analysis of rehabilitation therapy carried at infants'home.

3.6 Caretoy early intervention in infants with Down syndrome: a feasibility study

Down Syndrome is characterized by an alteration in developmental process which involves motor, language, cognitive, self-care and personal-social dimensions. Early intervention programs, focused on stimulation of developmental skills have shown some positive effects on development and outcomes. CareToy system, a smart tool for early intervention, has been recently developed and validated in preterm infants with positive short-term effects on motor and visual development. The main aim of this pilot study was to assess the feasibility of CareToy early intervention in a group of infants with Down Syndrome, secondarily it is aimed to evaluate the effects of the CareToy in promoting neurodevelopment in infants with Down syndrome.

3.6.1 Materials and methods

Participants

In this pilot study, the participation was proposed to families of eligible infant already followed by the clinical staff of IRCCS Fondazione Stella Maris (Calambrone, Pisa; Italy).

The Inclusion criteria were the following: 1) confirmed genetic diagnosis of DS, 2) age at recruitment between 3 and 9 months and achieved predefined cut-off score in gross motor ability derived from Ages & Stages Questionnaire Third Edition (ASQ-3; Squires & Bricker, 2009); that means, in general, that infants had yet acquired an initial head control but had not already acquired the trunk control. More in detail, ASQ-3 scores had to be as follow: i) 4 months form (from 3 months to 4 months 30 days) score ≥ 10; ii) 6 months form (from 5 months to 6 months 30 days) score ≥ 5 - < 50; iii) 8 months form (from 7 months to 8 months 30 days) score ≥ 10 - < 30.

The exclusion criteria were 1) severe sensory deficits (blindness or deafness); ii) progressive neurological disorder and iii) severe non neurological malformation or other medical conditions (e.g. percutaneous endoscopic gastrostomy or bowel derivation, recent cardiac surgery).

Detailed information about the study were provided to the families explaining that the allocation to CT or SC group was open and based on their availability to come to the clinical center either at least 3 times per week for 5 weeks to be allocated in the CT group or only 2 times at distance of 5 weeks to be allocated in the SC group, If parental written consent was obtained, infants were enrolled. Clinical data (demographic, medical and developmental data) were collected at the time of enrolment. At the beginning of the study period (baseline, T0) and after 5 weeks (T1), all infants were assessed with standardized clinical tools by clinicians and therapists blind to group allocation. The study was approved by Pediatric Ethics Committee of the Tuscan Region.

CareToy system

The CareToy System, as described above, is a biomechatronic system with a tele-rehabilitation architecture that allows to use the system at home with remote monitoring by the rehabilitation staff. Briefly, it is inspired to common baby gyms and it is composed by several modules: (i) a kit of sensorized toys designed to promote different grasping abilities, (ii) two interactive walls with lights, bottons and speakers, iii) a large monitor (screen wall), iv) a belt wall with a sensorized pillow, (v) an arch with lights, (vi) four cameras, (vii) a sensorized mat and (viii) three wearable sensors (two bracelets and chest strap). The combined activation of the modules and the feedbacks setting allow to provide a customized and active intervention, highly flexible, home-based and family-centered. In fact, during the CT activities, the infant is encouraged to actively perform specific goal-directed activities, aimed to promote postural control, reaching, grasping, visual attention and orientation in a highly variable way and at various level of complexity, at home and with his parents. The

CareToy is connected to a laptop which receive and send data from/to the clinical center (tele-rehabilitation module) and parents have a user manual (both printed and in the software) for the management of the training. Sensors are able to acquire and process data about the infant's behavior while he is playing (Donati et al; 2013; Rihar et al; 2014). A post processing elaboration creates an organized and readable report about infant's behavior for the clinical staff.

Even if the CT system is designed to be used at home, in this pilot study the system was installed and used in clinical environment. However, it was mainly managed by the parents under the supervision of a child therapist.

CareToy intervention

Clinical staff created a wide library of tailored and multi-axial goal-directed activities, called scenarios, projected on the basis of the most common developmental needs of infants and highly customizable not only in terms of active modules (e.g. one toy on a wall, two toys on the arch, the screen etc.) but also for other little features, as activity length, volume regulation, lights colors, videos and pictures presented on the screen.

CareToy training consisted in a set of scenarios, selected by the clinical staff on the basis of each infant's needs and capabilities. Each scenario lasts from 2 to 10 minutes and has a main goal and several sub-goals. CareToy training could promote many motor abilities such as postural control, rolling, pivoting, reaching, grasping, visual fixation and eye-hand coordination. The training is designed with a high degree of increasing variability and complexity. The management of CareToy intervention followed in general the guidelines described in the Sgandurra (et al 2016). According to them, first training days were planned on the basis of baseline assessment. The training was proposed for 5 weeks, at least 3 times per week and for 30-45 minutes daily to reach a minimum of 15 days and a range between 7.5 and 11.25 hours of total training.

Standard Care

Standard Care consisted in follow up visits, i.e. a monthly visit in which multidisciplinary staff of IRCCS Fondazione Stella Maris provides personalized advices to support parenting and to promote play activities and neurodevelopment of infant. In addition, if the infant underwent to specific interventions (i.e. physical therapy), the number of sessions per week was registered through a diary.

Outcome measures

Feasibility measures

CT training data and Scenario feasibility

The data of CT intervention in terms of differences between planned and executed days and minutes of training per day and total training were computed to assess the general feasibility of CT training. In addition, at the end of each scenario caregivers were asked to fill in the CT questionnaire about infant's participation (participation for: i) more than 75%, ii) between 50 and 75%, iii) between 25 and 50% or iv) less than 25%) and mood (the infant during the activity was mainly: i) happy, ii) neutral or iii) angry), to record the acceptability and feasibility of each proposed goal-directed activities and infant behavior.

Parents' point of view about usability and acceptability of CT training

A semi-structured interview, using the CareToy Questionnaire Parent-infant Experience (CQPE) (Inguaggiato et al; 2015), has been done with parents of children allocated to the CT group. The CQPE questionnaire is composed by 68 questions grouped into 9 macro-categories, aimed to explore different domains: 1) Parent's expectations regarding the CT training ; 2) Parent's opinion on infant's skill before the training 3) General opinions about the CareToy System, 4) Management of the system, 5) Time spent in the training, 6) Parent's opinion on changes on infant's skills after the training, 7) Parent' s participation during the activities, 8) Infant' s participation during the activities

and 9) Parent-infant interactions and home environment after training with CareToy.

Outcome measures

Motor assessment

According to our previous studies (Sgandurra et al; 2016), the primary outcome measure was the Infant Motor Profile (IMP) that is a new video-based and qualitative tool to assess the quality of motor performance and motor behavior in infants aged 3-18 months. It consists of 80 items that explore the child's motor abilities and behavior (Heineman et al., 2008, 2011, 2013). Moreover, in order to evaluate progresses in gross-motor abilities, we have used also Alberta Infant Motor Scale (AIMS): a standardized and widely used scale for infants from term up to 18 months of age, which allows to identify motor delays and/or abnormal motor development (Darrah et al., 1998). Both scales were carried out at T0 and T1.

Visual assessment

Teller Acuity Cards were used to evaluate the development of visual acuity. It is a behavioral validated test, for infant and children, based on forced-choice preferential looking and operant preferential looking stimuli with an observer's subjective assessment of an infant's looking behavior. The infant has to look to a series of grey cards that contain grating targets of various spatial frequencies. The examiner observes eye movements and behavior of the infant and judges whether the infant can or cannot see the grating on each card in the series. Acuity is estimated as the highest spatial frequency that the observer judges the infant to be able to see (McDonald et al; 1985; Teller 1990; Teller et al, 1986). It evaluates the development of visual acuity and it has been used in many studies for diagnostic purposes and to measure the results of early intervention. In the present study it was carried out during the assessment at T0-T1.

Data analysis

For each infant allocated in the CT group total days of training, mean of planned and executed training (in minutes) per day and mean of total planned and executed training (in hours) were calculated. In addition, a percentage ratio between the total of executed and of the planned CareToy training was calculated in order to assess the compliance to CareToy intervention.

Means and standard deviations (SD) of the infant characteristics and baseline measures for both groups were calculated. Mean delta changes in the primary and the secondary outcome measures were calculated between baseline (T0) and after the intervention period (T1). Potential trends due to intervention were qualitatively discussed in the results section. In relation to the small group of infants and to the only exploratory aim of the study, statistical analysis was not carried out.

3.6.2 Results

Study sample and characteristics

13 infants with DS were assessed for eligibility. 2 infants were excluded because they didn't meet inclusion criteria (gastro-intestinal and cardiac malformation with severe medical complications) and 1 family declined to participate. 10 infants were recruited between December 2015 and May 2016. All the infants who families requested to be allocated in the CT groups lived near the IRCCS Fondazione Stella Maris, with exception of one family (infant #3) who lived far (e.g. 70 kilometers).

5 infants (mean age months 6.64 ± 1.99, age range: 4.15–9.5) were allocated to the CareToy group and 5 (mean age months 5.6 ± 1.43; age range 4-7.5) to Standard Care group. The baseline characteristics of the participants are reported in Table 1. In the CT group none of the infants received any additional special rehabilitation sessions; while in the SC Group one of the five infants received physical therapy sessions (two times per week).

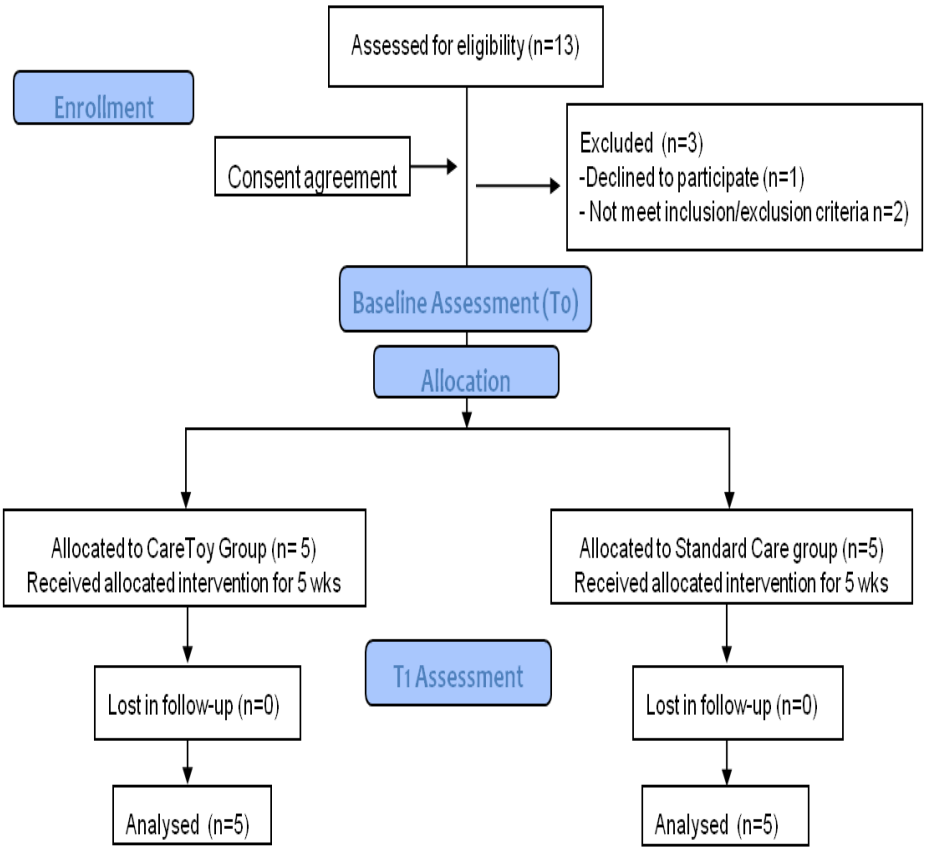


Figure 1: Consort Diagram of Pilot CareToy study in Down Syndrome

Table 1: Main characteristics of the CareToy and Standard Care groups at baseline (T0).

Characteristic	CareToy group (n=5)	Standard Care group (n=5)
Gender M/F	1/4	4/1
ASQ-3 Gross Motor Area		
Form 4 (median)	n= 2/5 (50)	n= 4/5 (50)
Form 6 (median)	n= 3/5 (20)	n= 1/5
Age at T0	6.64 (1.99)	5.6 (1.43)
Infant Motor Profile, mean (SD)		
Total	69.32 (6.02)	66.76 (1.98)
Performance	48.83 (10.2)	45.92 (5.93)
Variation	62.31 (4.48)	59.50 (0.25)
Fluency	73.33 (3.72)	75.00 (0.00)
Symmetry	96.19 (8.51)	90.95 (6.10)
AIMS Total, mean (SD)	12.60 (4.72)	10.2 (3.11)
Prone	4.40 (2.07)	3.80 (1.48)
Supine	4.60 (1.94)	4.00 (1.41)
Sit	2.60 (2.5)	1.40 (.89)
Standing	1	1
Teller Acuity Card (cy/degree), mean (SD)	5.82 (2.72)	4.76 (1.58)

Abbreviations: n, number; SD, standard deviation; M, male; F, female; ASQ-3, Ages & Stages Questionnaires; AIMS, Alberta Infant Motor Scales.

CT Training data

The 5 infants allocated in the CTgroup performed CT intervention for 5 weeks (daily planned mean in minute: 39.2 ± 4,02). The CT sessions were proposed at least 3 times per week; trainings were mainly performed in the morning. The training was managed mainly by the parents. A child therapist was present during the whole duration of the CT training, she observed and monitored the intervention in terms of system functioning, infants' behavior and parents'

ability in using the system, providing only minimum facilities to families (e.g. cleaning, turn on-off the system).

4/5 infants overcame the minimum of the 15 training days suggested, while one of infants (#3) performed only 9 days (Table 2). The time of executed training per day ranged from 25 to 39 minutes and total time of the training ranged from 5.16 to 17.54. The ratio between planned and executed training ranged from 66.31 to 99.62 %.

Table 2: Mean duration (± SD) of the CareToy training in days; mean daily planned and executed training in minutes, total executed training in hours and the ratio between the total of the executed and the planned training in percentage.

User	Training days (n°)	Mean planned per day (min)	Mean of executed per day (min)	Total executed (hrs)	Ratio (%) planned/executed training
#1	23	40	27	10,23	66,31
#2	21	39	30	10,32	74,85
#3	9	45	34	5,16	47,76
#4	27	39	39	17,54	99,62
#5	18	33	25	7,06	76,67
mean ± SD	**19,4 ± 6,84**	**39.2 ± 4,02**	**31 ± 5,61**	**10,06 ± 4,71**	**73,04 ± 18,75**

Abbreviations: SD: standard deviation, N: number, min: minutes; h: hours

CT scenarios feasibility

A mean of 30% of scenarios were participated for more than 75%, 65% of scenarios for 50–75%, and only 5% scenarios for 25–50%. A mean of 90% of scenarios were performed with good grade of acceptance of the system and playing activity (happy mood of infant), 9% were performed with a medium grade of acceptance (neutral mood infant) and only 1% were not enjoyed by the infants (angry during the activities).

CPQE

4 of 5 families involved in the CT training were interviewed. Their expectations (Domain 1) were that the CT training could mainly improve the

psychomotor development of their infant. All families judged CareToy experience as positive (Domain 3) mainly because the CT intervention allowed to improve the attention, gross motor (i.e. rolling) and hand abilities (i.e. grasping) of their infants (Domain 2 and 6). The system management (Domain 4) was evaluated as ‘‘rather simple’’. The training activities were considered as ‘‘well personalized’’ (3/4) or ‘‘enough personalized’’ (1/4) for their infants’ needs. The infant's participation (Domain 8) was deemed, by all parents, as generally good by saying (4/4) that "the child was quiet for most of the time" and that (4/4) "he/she was interested in most of the activities". The reported activities that infants liked more where those that mainly involved the use of the screen or toys (especially when they were attached to the arch); on the other hand scenarios involving the lights (on the walls or on the arch) were deemed as less interesting. Concerning the participation of parents (Domain 7) during the training, CPQE data indicated that all parents felt free to interact and participate during their infant's play activities during the CT training. Finally, for the Domain 9 parents (3/4) stated that, after training, they have changed the way they play with their infant "in large part"; in one case they changed "completely" (1/4).

Clinical results

Motor outcome measure

After the intervention period the delta changes (T1-T0, mean ± SD) of IMP Total score were higher in the CT group (3.05 ± 4.52) than in the SC group (2.13 ± 1.62; Table 3). Moreover in the IMP Performance domain after the intervention period an higher delta changes difference was found between the two groups (CT vs SC; 7.16 ± 5.35 vs 3.90 ± 4.77). After the intervention period a difference was found also in the AIMS Total in which CT group (4.40 ± 3.78) reached an higher delta changes compared to the SC group (2.6 1.51; Table 3).

Visual outcome measures

After the intervention period the delta changes (T1-T0, mean ± SD) in the visual acuity increasedin the CareToy group compared to Standard Care (1.50 ± 1.62 vs 0.80 ± 1.39); Table 3).

Table 3: Mean of delta changes (T1-T0) of the two groups

Characteristic	CareToy group (n=5)	Standard Care group (n=5)
Infant Motor Profile, mean (SD)		
Total	3.05 (4.52)	2.13 (1.62)
Performance	7.16 (5.35)	3.90 (4.77)
Variation	2. 68 (5.77)	1.93 (2.28)
Fluency	0.00 (0.00)	0.00 (0.00)
Symmetry	0.00 (0.00)	2.48 (5.95)
AIMS Total, mean (SD)	4.40 (3.78)	2.6 (1.52)
Prone	1.60 (1.34)	0.80 (0.84)
Supine	0.80 (0.84)	1.60 (1.34)
Sit	1.60 (2.5)	0 (0)
Standing	0.40 (3.78)	0.20 (0.45)
Teller Acuity Card (cy/degree), mean (SD)	1.50 (1.62)	0.80 (1.39)

Abbreviations: n, number; SD, standard deviation; M, male; F, female; ASQ-3, Ages & Stages Questionnaires; AIMS, Alberta Infant Motor Scales.

3.6.3 Discussion30

This pilot study was mainly devoted to test the feasibility and the acceptability of the CT EI in a group of infants with DS; secondarily the study aimed to explore the trends in the effects of the CT intervention, compared to standard care, in promoting motor and visual development.

Feasibility of CareToy intervention as a tool for EI in Down Syndrome

The modularity of CareToy system and the wide variability of the scenarios library have allowed to plan, for each infant, a tailored and multi-axial training. In general scenarios already present in the library of the CT system demonstrated to be suitable and useful also for infants with DS.

In some occasions, scenarios were slightly adapted and this customization mainly involved the length. The global duration of training has been maintained but the single scenario duration demonstrated to need a change and this permitted to guarantee a high quantity of daily training in infants with DS. An interesting data about the training regarded also the variability: infants with DS were interested for much time to the same scenario and demonstrated to appreciate and need more repetitions of the same scenario for better pay attention to stimuli and organize their motor activity. Another small feature adapted to this population of infants was the requested pressure on toys for evocate the feedback, in fact the value was reduced. With these little adjustments of scenarios, the applied scenarios were suitable for infants' needs and, after the firsts sessions, the intervention of the therapist was never required. This means that the feasibility of the training is referred also to the management of the training by parents.

4/5 infants overcome the minimum number of training days requested in the study; only the infant #3 performed only 9 days of training and in total less than 50% of the training, this data could be related to high morbidity of the infant (frequent fever). However, to reach the cut off, families spent a lot of time and energy to do the intervention and this finding points out the advantage

of providing home-based intervention, especially in the early age of development, increasing the time of intervention and thus the intensity. Data obtained from CT questionnaire, filled in by the parents immediately after each CT scenario, reported encouraging data about the compliance of infants because the majority of scenarios were played and enjoyed by the infants and were participated for more than 50% of the duration.

CQPE data also highlighted that all the parents have considered the CareToy experience as positive, deem as ‘‘well personalized’’ or ‘‘enough personalized’’ the scenarios activities. All these data support the good acceptability and the feasibility of CareToy training also in infants with DS.

Thanks to the modularity of the CareToy scenario library and to the possibility to customize also minor features of scenarios, some modifications of activities were made at the beginning of the study but also during the experimental trial, in order to adapt the CareToy intervention also to infants with DS taking account their specific developmental needs. It is known that infants with DS often present cognitive delay and their time of response to stimuli or their initiative could be delayed, for this reason scenario's length has been increased, in order to give them more time for perceive the visual and/or audio stimuli and orientate to it, or to organize the motor activity. As in the previous project, scenarios' duration has been calibrated to each single infant.

Another characteristic of infants with DS is the global hypotonia, that could influence the postural development; based on this issue feedback events have been set at lower values, which means that the requested pressure for activate the response was lower than the one planned for the population of preterm infants.

As said before, the CareToy intervention was previously designed for preterm infants and we hypothesized that it could be feasible also for infants with DS. Some small modifications of the CT activities have been planned before the training and tested on going.

Effects of CareToy intervention

All the recruited infants performed all assessments and interventions according to their allocation. Regarding the characteristics of sample in general the baseline data (T0) were slightly different between the two groups (CareToy and Standard Care) mainly due to high variability confirmed by the wide SD (Table 1). In all outcome measures, the mean delta changes (T1-T0) showed promising results highlighting positive effects in the CareToy group compared to Standard Care one, in fact in experimental group we found higher changes than in the control group. This effect appeared more evident in the motor scales (both IMP and AIMS) than in visual assessment. Moreover, results could be, in part, limited by the used motor measures; in fact the IMP, used as a primary outcome measure and chosen to take the same approach of the previous study, is not a validated tool to assess motor development in DS. On the contrary the AIMS is a validated tool in the DS, but unlike to the IMP, it measures mainly aspects of gross-motor development and could not able to evaluate changes in short period as in the previous study with preterm infants. In general the small size of the sample and the large SD determined a wide variation in the results, so the results appear promising but not conclusive and need further confirmation with a more extensive sample and an higher quality study.

Limitations

This study has some limitations. In particular the sample in the study is small and the infants' allocation was related to the possibility of the family to attend the clinical center, so study lacked the randomization process lacks randomization. The home-based intervention and tele-rehabilitation approach can overcome this issue enlarging the access of infants to EI programs. Moreover some outcome measures need adjustments for the population in question and follow-up evaluations should be provided to evaluate medium and long-term effects. These data were out of our principal aim and there will be overcome in the future study with CareToy in infants with DS.

3.7 Final remarks

The importance of early identification of developmental disorders is raising also because there is an increasing number of rehabilitative approaches.
In the modern panorama of infants' assessment and rehabilitation, technology play a crucial role: it allows to quantitative measure and process data about infants' capabilities and this could represent in the future a new way to summarize infants' competences, creating a new database with normative data relative, for example, to manipulation capabilities and strength, standardizing and facilitating the assessment and the comparison between subjects. The presented results verify that the CareToy system can be suitably used for highly relevant assessment and stimulation purposes, moving both the assessment set-up and rehabilitative programs directly to families' home. A correlation with clinical assessment scores has been demonstrated and the proposed training activities and goals were showed as adequately designed to stimulate task-specific infant activity responses. Also, infants are assessed while they are doing what they usually do, that is play. It is expected that the proposed home environment-based stimulation and assessment system could, besides the ability of realistic toy play and natural environment, due to high accuracy, reliability, and objectivity of integrated sensor systems provide the opportunity i) to practice activities, more enjoyable than traditional therapy tasks, thereby encouraging training intensity through high trial repetitions and ii) to perform a family and child-centred approach, supporting them at implementation of the rehabilitation training.
The comparison of results will most definitely be interesting also for clinicians and could provide the possibility of further data analyses. Moreover, a comprehensive analysis of clinical and instrumental results could provide predictive indicators of responses to intervention, as well as correlation between the developmental changes and the amount of training with the CareToy system. Also, the RCT study has strengthened evidence that the CareToy system can provide effective home-based, individualized early

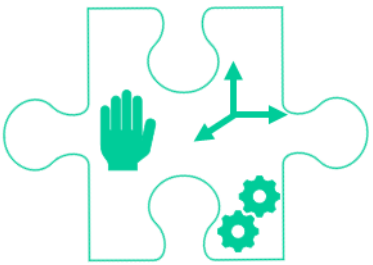
intervention, for sure in preterm infants without severe medical complications and it started to investigate if could promote the development also of infants with Down syndrome, stating that it has been confirmed its feasibility and suitability. CareToy intervention is a strong candidate for testing on a larger scale with the aim of achieving long-term benefits to children with increased risk of developing neurological impairments. Another important aspect of this new approach, in relation to its transferability to clinical practice, deals with cost of this technology, when ready for industrialization. Recently, an early-stage economic model (Manetti et al, 2015) for the evaluation of these aspects has been developed. Additionally, training trials on other populations, such as preterm and term infants at high risk of developing cerebral palsy due to congenital brain lesions or infants with developmental delay due to genetic abnormalities needs to be planned for additional identification of system suitability for providing assessment, stimulation, monitoring and training of infants in the first year of life.

Data reported in this Chapter are based on the following paper:

- Passetti G, Cecchi F, Baldoli I, Sgandurra G, **Beani E**, Cioni G, Laschi C, Dario P. *Sensorized toys for measuring manipulation capabilities of infants at home* Conf Proc IEEE Eng Med Biol Soc. 2015;2015:7390-3. doi: 10.1109/EMBC.2015.7320099
- Rihar A, Sgandurra G, **Beani E**, Cecchi F, Pašič J, Cioni G, Dario P, Mihelj M, Munih M *CareToy: Stimulation and Assessment of Preterm Infant's Activity Using a Novel Sensorized System.* Ann Biomed Eng. 2016 Dec;44(12):3593-3605. Epub 2016 Jun 10.
- Sgandurra G, Lorentzen J, Inguaggiato E, Bartalena L, **Beani E**, Cecchi F, Dario P, Giampietri M, Greisen G, Herskind A, Nielsen JB, Rossi G, Cioni G; CareToy Consortium. *A randomized clinical trial in preterm infants on the effects of a home-based early intervention with the 'CareToy System* PLoS One. 2017 Mar 22;12(3):e0173521 doi:10.1371/journal.pone.0173521. eCollection 2017
- Baldoli I, Cecchi F, Passetti G, Sgandurra G, Inguaggiato E, **Beani E**, Cioni G, Mihelj M, Laschi C, Dario P, *Sensorized toys for measuring the effect of home-based therapy on manipulation capabilities of infants,* submitted
- Inguaggiato E, Sgandurra G, Bargagna S, **Beani E**, Orlando M, Purpura G, Cecchi F, Dario P, Cioni G and CareToy Consortium *CareToy Early Intervention in infants with Down Syndrome: a feasibility study*, submitted

CHAPTER 4

QUANTITATIVE ASSESSMENT OF UPPER LIMB USE IN CHILDREN WITH TYPICAL DEVELOPMENT AND WITH NEURODEVELOPMENTAL DISORDERS BY INERTIAL SENSORS



4.1 Introduction

Upper limb function is extremely important because during the development guides infants in discovering the world and then it is fundamental for basic and instrumental activities of daily living, as self-care, work, leisure, and household activities, and for social communication. These activities include both global and fine-motor movements and involve the cooperation between upper limbs, mainly characterized by a distinction between right and left, since one of the hand conducts the action, whereas the contralateral helps in completing the motor task, playing a secondary role. For this reason, the two hands are commonly called "dominant" and "non-dominant", respectively. The presence of a more marked lateralization of the function to the right or left arm aims at the optimization of the bilateral collaboration between the two sides, referred to as "handedness" (Duroz, 2014).

The hand is an effective tool that is used in many different tasks of daily life. The usefulness of the hand is highly dependent on cognition, because we have to understand the value of using hands for a meaningful purpose. Use of the hand consists of various aspects, such as reaching, grasping, manipulation, unimanual and bimanual actions and grading of grip force.

Cerebral palsy is the most common cause of childhood chronic physical disability in Europe and in other industrialised societies. The incident rate is between 2 and 3 per 1000 live births and increases to 40–100 per 1000 live births among very premature and very low birth-weight babies which represent the population with highest rate of neurodevelopmental disorders. Unilateral Cerebral Palsy constitutes the most frequent form of cerebral palsy, about 38% of all affected children. Recent estimations of incidence and prevalence of cerebral palsy have shown a significant increase in unilateral cerebral palsy in Europe over the last years. The upper limb of children with unilateral cerebral palsy is generally more involved than the lower limb and the consequent

disability affects their participation, quality of life, independence at home, school and community (Cioni et al, 2010).

The assessment of upper limb function is a widely investigated aspect of the development and its importance is related to the need of the identification, description and rehabilitation of the impaired arm and hand for routine bimanual activities and to achieve functional independence.

Beside traditional clinical tools (seen in chapter 1), nowadays different technologies are used to investigate hand function. Some of them include electrophysiologic assessment, e.g. Transcranial Magnetic Stimulation (TMS) and kinematic analysis: inertial sensors are now being validated as a new rehabilitation instrument. The kinematic analysis is a powerful instrument for a quantitative assessment in movements analysis. It provides non-invasive three dimensional motor analysis data such as kinematics and kinetics. It can also give a quantitative evaluation of muscle activity if associated to electromyography. Unlike the lower limb where multiple gait laboratories value the use of the limb, the use of the upper limb is more complex and variable so the assessment and interpretation of data is really challenging. Inertial sensors are used in a considerable number of research fields in medicine to detect different sorts of movements and there are in literature an increasing number of studies focused on infants and children.

After and together to assessment, also rehabilitative approaches are improving. Concerning the rehabilitation models commonly used for hemiplegia, there are fewer available therapies and many studies are under way to find new strategies to improve the use and function of the upper limb so as to increase both hand manipulation and the quality of life of the patients (Ghai et al, 2013). In a recent review (Sakzewski et al, 2009) the authors identified four main interventions: intramuscular botulinum toxin A (BoNT-A) combined with upper-limb training; constraint-induced movement therapy (CIMT); hand-arm bimanual intensive training (HABIT) and neurodevelopmental therapy. They demonstrated that no one treatment approach seems to be superior; however, injections of BoNT-A provide a supplementary benefit to a variety of upper

limb-training approaches. Both CIMT and HABIT rely on upper limb intensive training, either unimanual or bimanual, respectively. They are underpinned by theories of motor learning and neuroplasticity that describe a correlation between improved motor function and the use of "massive" or "repetitive" practice.

Recent neurophysiological evidence on motor control and learning has disclosed some interesting opportunities for rehabilitation. Discovery of mirror neuron system and its role in motor learning has recently revolutionized the field of upper limb rehabilitation. A novel paradigm defined 'Action–Observation Training', based on observation of meaningful actions followed by execution, has been introduced as a rehabilitation therapy for the first time in 2007 (Ertelt et al, 2007). Then it has been recently proposed and used with promising results in clinical studies, mainly on adult stroke patients (a review is presented below). Moreover, recent studies carried out on children with cerebral palsy indicate positive effects on upper limb function, in particular, in Sgandurra et al (2011 and 2013) recently carried out a clinical study called UP-CAT (UPper limb Children Action observation Training) on children with unilateral cerebral palsy based on AOT, providing evidence of its efficacy in improving upper limb activity performance in daily activities.

Starting from these findings, an interesting development of this approach could be to add quantitative measures of upper limb activity and to test if also a home-based training could be suitable for this type of training. In this framework, Information and Communication Technologies (ICT), as wearable sensors mentioned above, could represent a viable option in providing AOT programs at home, in a user-friendly, playful and rehabilitative setting; in fact, ICT allows to quantitatively measure and monitor daily upper limb activity and to detect changes of upper limb activities during and after the AOT program.

4.2 Assessment of upper limb use in children with typical development and with neurodevelopmental disorders by inertial sensors: a systematic review

Understanding the development of bimanual upper limb (UL) activities both in typical and atypical conditions in children is very important. New technologies, such as wearable sensors, are increasing their use, more on adults than in children. The main aim of this section is to explore the extent of upper limb activity of both typically developed (TD) children and children with neurodevelopmental disorders (NDDs) that can be reliably recorded and compared with each other, using a combination of multiple wearing inertial sensors, both in a standardized and a natural environment.

4.2.1 Methods

Research Questions

This systematic review aimed to address the following specific questions: 1) What are the clinically relevant research applications of UL inertial sensors, both in a standardized and a natural environment? 2) How valid and reliable are at least two bilateral arm-hand worn sensors, with regard to the number and the specific placement of the devices, in order to i) describe UL motor capacity and performance in children and ii) investigate the UL bilateral asymmetry using accelerometers; 3) Can the assessment be generalized across different UL motor abilities and health conditions? 4) What types of accelerometers have been used and how should data be collected and analysed? Are the measurements obtained related to clinical assessment?

The criteria for including the studies in this review concerned the types of participants, the forms of assessment and the outcome measures. The present review has been registered at Prospero CRD42016033687.

Types of Participants

Studies in this review were included if paediatric subjects were evaluated. The established age range includes participants between 0 to 20 years. Twofold studies assessing both adults and children were considered, omitting the adult sample data and analysis from our study, and focusing on the younger participant results only. Studies including both healthy and/or subjects with disabilities were included.

Forms of Assessment

According to Quality Assessment of Diagnostic Accuracy Studies, version 2 (QUADAS-2) guidelines (Whiting, P., Rutjes, A. W., Reitsma, J. B., Bossuyt, P. M., & Kleijnen, J. 2003), in this review index test and reference standard are considered. In particular, inertial sensors are the index test, namely the tool which is being evaluated against a reference standard, when referred, to evaluate their accuracy. Clinical tests are considered as reference standard. For the index test, because of various models of inertial sensors available, the type of devices was not a limitation for inclusion criteria.

In addition, all participants wore at least 2 accelerometers on both of the UL at the same time. All the studies reporting comparisons between activity of both arms, recorded by inertial sensors, were included in the review. As concerning the reference standard, it had to have the capability to assess and to quantify the amount of physical activity of the two ULs, both as separated unilateral data and compared to each other.

Outcome Measures

Data extraction of outcome measures of interest included the amount (raw count) and duration (percentage of the recording period) of movements and the intensity of physical activity (acceleration vector magnitude namely the 3D summing vector of the three axes), performed by the ULs to detect the potential bilateral asymmetry.

Criteria for studies inclusion

Searches were conducted between September 2015 and May 2017 (last search May 16th, 2017) using the following databases: Pubmed, Web of Science and EBSCO (CINHAL®Complete), establishing the English language. The following terms were utilized for searches: (actigraph* OR acceleromet* OR inertial) AND ("hand dominance" OR handedness OR "upper limb" OR laterality OR asymmetr* OR "upper extremit*" OR "movement ratio" OR "movement pattern" OR wrist OR hand OR arm) AND (child* OR adolescent OR teen*). Review publications were included. The two first authors (IB, MM) and co-authors (GS, EI, EB, FC) conducted all searches.

Eligibility

The following criteria were required for inclusion: i) simultaneous data collection of at least 2 inertial sensors, on each UL. Specific placement of the inertial sensors could vary across studies (wrist, finger, forearm, etc); ii) report of the relationship between the recorded data of the two extremities, reported as asymmetric coefficient (or similar), comparisons, algorithms, ratios, etc.; iii) human participants aged 0-20 years. Exclusion criteria included: i) absence of simultaneous data recording between the two ULs; ii) use of non-human subjects or simulation-based data; iii) failure to report data counts from inertial sensors and/or their correlative relationship; iv) use of out-of-center measurement device to determine physical activity other than inertial sensors (frequency counters, electronic thermometers, Molecular Electronic Transducers (MET), direct observations, questionnaire).

All articles were screened for inclusion by the authors (GS, IB, MM, EI, FC), unblinded to manuscript authorship.

Data extraction

Authors extracted all data independently (unblinded). Extracted data included both clinical and technical details of the included studies.

Clinical Data

The clinical data investigated are the following: author/journal, year of publication, type of study, potential different measurements used in the study other than inertial sensors, number and demographics (ages, sex, BMI, height, weight, waist circumference, co-morbid disorders and their potential reported details, etc.) of subjects, setting of the study, inclusion/exclusion criteria of participants, aims, limitations and conclusions of the study.

Technical Data

The technical data analysed are the following: make/model and number of inertial sensor, placement of device, time of wearing of the device, sample acceleration (Hz), accelerometer activity measures, data cleaning, statistical approach used for data analysis, threshold to assess the intensity of arm movement, threshold as cut-off frequency of filter applied on raw data, main features for accelerometer data comparison, differences between the two hands.

Data for TD and NDDs participants were reported separately.

Assessment Study Quality

Study quality was assessed (unblinded) by IB and GS using the standards of the QUADAS-2. Specific questions were added in four domains by the authors, to bring this quality assessment tool into line with the main review question. These included the Patient Selection Domain, where the question "Are the data for paediatric sample well defined?" was added, in the Applicability Concerns Section. In terms of Risk of Bias, both the Index Text and the Reference Standard Domains had been extended, with the following two questions, deemed fundamental: 1) Has the index test been used, in order to detect the presence of differences in the physical activity of the two ULs? 2) Have the data of the index test been clearly and correctly reported in order to match the review question?

Regarding the Flow and Timing Domain, all the first three standard questions were refined, in order to customize the reference standard to the review question.

Ratings of each study using QUADAS-2 are presented in the Fig. 2, 3 and 4 as a resource/reference to the reader, but were not used in the weighting of quantitative data.

4.2.2 Results

Study Inclusion and Assessment

The Preferred Reporting Items for Systematic Reviews (PRISMA) flow diagram is presented in Fig. 1. A total of 1127 articles were identified by the initial search criteria. After repeats and duplicates were removed, 752 possible records were identified, which were subsequently screened for inclusion/exclusion, according to the title and the abstract. After revision, 92 full articles were retrieved for closer consideration and 72 of them did not meet our criteria for qualitative review. Reasons for exclusion were grouped into four main categories: 1) Unsuitable placement and/or number of sensors: n=27; 2) No inertial sensors (i.e.: inertial eigenvectors, dynamometers, potentiometers): n=3; 3) Age (i.e.: adult): n=17; 4) Unsuitable data analysis (i.e.: absence of evaluations regarding the potential bilateral asymmetry existing between the two ULs): n=19; 5) Other reasons (i.e.: descriptive texts, full article not available): n=6.

Amongst them, one dissertation was excluded as an unpublished work. Since it however met all the required criteria of the review, it was later included in the analysis, as additional record, identified through other sources (Davila E. M. 2011).

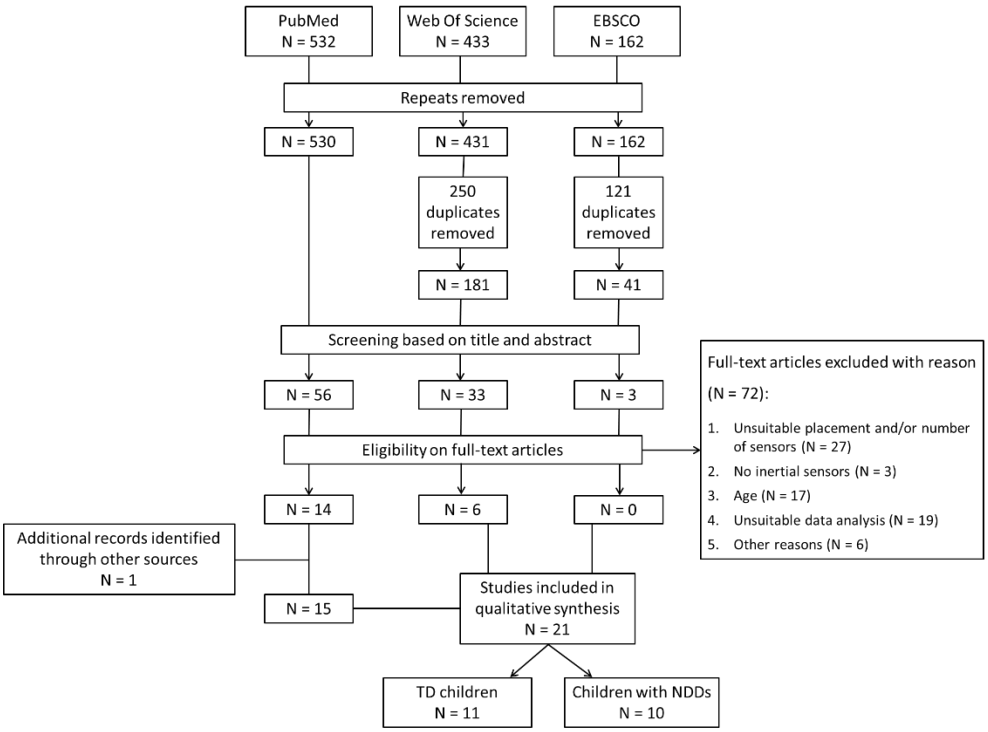


Figure 1

In conclusion, there were 21 articles remaining that met the full inclusion criteria (Sokal, B., Uswatte, G., Vogtle, L., Byrom, E., & Barman, J. 2015; Lemmens, R. J. et al. 2015; Zoccolillo, L. et al. 2015; Phillips, L. R., Parfitt, G., & Rowlands, A. V. 2013; Gordon, A. M., Schneider, J. A., Chinnan, A., & Charles, J. R. 2007; Floyd, A. G. et al. 2007; Sadeh, A., Sharkey, M., & Carskadon, M. A. 1994; Birmingham, A. T., Wharrad, H. J., & Williams, E. J. 1985; Bergamini E. et al. 2015; Deutsch, K. M., & Newell, K. M. 2006; MacArthur, B., Coe, D., Sweet, A., & Raynor, H. 2014; Strohrmann, C. et al. 2013; Davila E. M. 2011; O'Neil, M. E et al. 2016; Kaneko, M., Yamashita, Y., Inomoto, O., & Iramina, K. 2015; Kaneko, M., Yamashita, Y., & Iramina, K. 2016; Le Moing A. G., et al. 2016; Graves, L. E., Ridgers, N. D., & Stratton, G. 2008; Dadashi, F., Millet, G. P., & Aminian, K. 2016; Mackintosh, K. A., Montoye, A. H. K., Pfeiffer, K. A., & McNarry, M. A. 2016; Coker-Bolt, P. et al. 2017). The number of eligible articles increased exponentially with time.

Of these 21 studies, most were conducted in the United States (34%), followed by the United Kingdom (19%), Italy, Switzerland and Japan (9% each), and finally France, Israel, Australia and Netherlands (5% each). Study quality was rated using QUADAS-2 ratings.

Participants Characteristics

Participants characteristics are reported in Table 1 and 2. Considering both the two groups of children (TD and NDDs), they were aged between 2,2 to 20 years. The gender distribution was generally equally balanced. In 11 articles TD children were enrolled, both volunteers and randomly selected from schools, fitness classes and from broad geographical areas (Dadashi, F., et al. 2016; Mackintosh, K. A., et al. 2016; Birmingham, A. T., et al. 1985; Graves, L. E., et al. 2008; Lemmens, R. J. et al. 2015; Phillips, L. R., et al. 2013; Sadeh, A., et al. 1994; Deutsch, K. M., & Newell, K. M. 2006; MacArthur, B., et al. 2014, Davila, E. M. 2011; Kaneko, M., et al. 2015). The other 10 articles included children with NDDs, whose limb motor ability, especially the ULs, were impaired (Zoccolillo, L. et al. 2015; Sokal, B., et al. 2015; Gordon, A. M., et al. 2007; Floyd, A. G. et al. 2007; Bergamini, E. et al. 2015; Strohrmann, C. et al. 2013; O'Neil, M. E et al. 2016; Kaneko, M., et al. 2016; Le Moing A. G., et al. 2016; Coker-Bolt, P. et al. 2017). The type of NDD are: i) right or left hemiparesis subsequent to CP or to postnatal stroke (O'Neil, M. E et al. 2016; Sokal, B., et al. 2015; Zoccolillo, L. et al. 2015; Gordon, A. M., et al. 2007; Strohrmann, C. et al. 2013; Coker-Bolt, P. et al. 2017), showing a mild to moderate motor impairment; ii) ADHD (Kaneko, M., et al. 2016); iii) Duchenne Muscolar Dystrophy (Le Moing A. G., et al. 2016), including only non-ambulant patients; iv) Niemann-Pick disease (NP-C) (Floyd, A. G. et al. 2007) with moderate ambulatory impairment and moderate to severe disability in one or more functional system (Ataxia, Dystonia, dysmetria, myoclonic jerks, tremor, peripheral neuropaty, etc.); v) paraplegia (Bergamini E. et al. 2015), due to several diseases such as myelomeningocele, poliomyelitis, diplegic cerebral palsy, bone cancer.

Table 1: Clinical Data in Typically Developing Children.

Author	Study type	Aims	Setting	Sample Size	Mean Age (yrs)	Inclusion Criteria	Exclusion Criteria	Reference Standard
Birmingham A. T. et al. ***(1985)***	Survey	To evaluate the variation of tremor frequency and amplitude in relation to the age	Laboratory	109	n=22 (7-9), n=28 (9-11), n=24 (11-13), n=22 (13-15), n=13 (16-18)	To be attending specific schools or fit classes	NA	NA
Avi Sadeh et al. ***(1994)*** **[STUDY 1]**	Lab-based validation and calibration study	To develop a new sleep-wake scoring algorithm	Laboratory	16	13.8±1.9	Volunteers	NA	NA
Deutsch K. M. et al. ***(2006)***	Observational study	To investigate the mechanical and neural components of postural finger tremor	Laboratory	39	n=20 (6.4±3), n=19 (10.5±0.3), n= 21 (20.8±1.4)	Volunteers	Neurological disorders, influencing tremor	NA
Graves L.E.S. et al. ***(2008)***	Observational study	To examine the contribution of the upper limb and total body movement to adolescents' energy expenditure whilst playing videogames	Laboratory	13	11-17	Good health picture	NA	NA

Davila E. M. ***(2011)***	Observational study	To evaluate the influence of wearing AMs on the D vs ND wrists on measurements of free living PA	Laboratory + outpatient	20	12-17	1) volunteer participants from Bozeman, Montana, 2) 12-17 yrs	NA	NA
Phillips L. R. S. et al. ***(2012)***	Lab-based validation and calibration study	To develop physical activity intensity cut-points for use with GENEA accelerometer	Laboratory	44	10.9 ± 1.9	NA	NA	NA
MacArthur B. et al. ***(2014)***	2x2 mixed design with random allocation.	To measure percentage of time engaged in MVPA and estimated EE with accelerometry in playing AVG vs unstructured OP	ELC playground, ELC room	16 (8 OP vs. 8 AVG)	OP= 6.6±0.7, AVG= 6.3±0.9	1) good health, 2) healthy weight (BMI percentile= 5-85), 3) no limit for physical activity.	Grass allergies. Skin sensitivity to light. Failure to complete all session within a 3-week period.	NA
Lemmens R. J. M. et al. ***(2015)***	Cross sectional study	To evaluate the reliability of arm-hand tasks accelerometer records	Laboratory	32	n=16 children (8.5±1.7), n=16 (14.6±1.5)	Volunteers	Motor problems with arm, hand or shoulder	NA
Kaneko M. et al. ***(2015)***	Observational Study	To quantify age-appropriate developmental changes of SNS	Laboratory	233	4-12	Student at the Fukuoka Municipal Elementary School	NA	NA
Dadashi F. et al. ***(2015)*** **[GROUP 2]**	Observational study	To characterise front-crawl swimming skills	Outdoor pool	9	16.0 ±1.8	Recreational swimmers	NA	NA

Mackintosh K.A. et al. ***(2016)***	Observational study	To validate and compare ANNs	Laboratory, semi-structured setting	27	10.8 ±1.0	Volunteers recruited via a local primary school. The children attended the the laboratory only if: 1) rested state, 2) at least 2h postprandial, 3) strenous exercise and coffeine avoided in the previous 24h	NA	NA

ANNs: Artificial Neural Networks, AVG : Active Videogames, EE: Energy Expenditure, ELC: Early Learning Center, OP: Outdoor Play, NA: not available

Table 2: Clinical Data in Children with Neurodevelopmental Disorders

Author	Study type	Aims	Setting	Sample Size	Disease	Mean Age (yrs)	Inclusion Criteria	Exclusion Criteria	Reference Standard
Floyd A. G. et al. *(2007)*	Multy-centered study	To analyze the UL motor physiology	Laboratory	15	NP-C	25 ± 10	1) => 12 years of age, 2) confirmed diagnosis of NP-C by abnormal cholesterol esterification and abnormal filipin staining	1) concurrent enrolment in other clinical trials, 2) drugs or diet supplements, interfering with digestive absorption of study medication, 3) significant history of gastrointestinal disorders, HIV or hepatitis, 4) not comply with study procedures	EDSS
Gordon A. M. et al. *(2007)*	Single-blinded randomized control study	To examine the efficacy of the HABIT	Summer camp	20 (10 HABIT vs 10 CG)	UCP	total sample = 9.6±6.0, HABIT = 4.5-13.7, CG= 3.9-10.6	1) ability to extend the wrist>20° and the fingers at the metacarpophalangeal joints>10° from full flexion, 2) JTHF: >50% difference between the involved and the non-involved hand, 3) ability to lift the involved arm>6 inches, 4) BBIT= mean score +/-<1DS	1) health problems unassociated with CP, 2) current/untreated seizures, 3) visual problems interfering with the intervention or testing, 4) MAS> 3.5, 5) orthopaedic surgery on the involved upper extremity, 6) dorsal rhizotomy, 7) botox therapy in the UL in the prior 6 months or within the period of study, 8) intrathecal baclofen.	AHA, BOT-2, CFUS, JTHF

Strohrmann C. et al. *(2013)*	Longitudinal study	i) to monitor children activities in daily life, ii) to evaluate the use of body worn sensors for motor assessment in children	Laboratory	4	CP (2), acquired stroke (2)	10.5 ± 2,12	1) neurological diagnosis leading to stationary stay, 2) age= 5-18 years, 3) cognitive ability to understand the aim of the tasks	NA	Motor Capacity Assessment
Zoccolillo L. et al. *(2015)*	cross-sectional experimental quantitative study*	i) to monitor physical activity during VGT vs conventional therapy, ii) to quantify if VGT enhances number of movements	Outpatient + inpatient	8	UCP	6,6 ± 1,4	1) UCP, 2) 4-14 yrs, 3) GMFCS: I-IV, 4) any Xbox with Kinect at home.	1) IQ<35, 2) severe comorbidities, 3) incapacity to stand, even with an external support.	QUEST, ABILHAND-kids

Sokal B. et al. ***(2015)***	Cross sectional, observational design	i) to evaluate the UL activity, ii) to compare the use of the affected arm between children and adult with hemiplegia	Not reported	28	UCP	3.9 ± 1.7	NA	1) serious or recurring medical complications, 2) spasticity medication within the last 3 months, 3) previous paediatric CIMT, 4) fixed contractures in the affected-arm, 5) invalid accelerometer records (insufficient time, only 1 wrist, unrealistic records, malfunction)	PMAL-R, PAFT
Bergamini E. ***(2014)***	Three experimental sessions	i) to identify a biomechanical performance indicators of wheelchair propulsion, ii) develop and assess the efficacy of a specific training program	Basketball court	12 (6 EG vs 6 CG)	paraplegia (4), myelomeningocele (3), poliomyelitis (2), spastic diplegia (1), below-knee amputation (1), knee arthroprothesis (1)	total sample = 17.1 ± 2.7, EG= 13-20, CG= 12-20	At least two years of previous wheelchair basketball experience.	Medical contraindications	NA

Kaneko M. et al. *(2016)*	Observational study	i) to establish a quantitative evaluation system of soft neurological signs	Laboratory	33	ADHD	7-11	1) patients of the Kurume University Hospital, 2) positive DSM-IV criteria for ADHD diagnosis, 3) WISC-III>70	NA	NA
Le Moing A.G. et al. *(2016)*	Observational study	i) to highlight the feasibility of quantifying the range of upper limb movements	Laboratory	7	DMD	18.5+/-5.5	1) patients of the Institute of Myology, 2) age> 10 years old, 3) non-ambulant, 4) able to sit for at least 3 hours in the wheelchair	1) cognitive impairment, 2) occurrence of neurological/inflammatory/infectious/endocrine/acute orthopaedic disease in the precious month, 3) scheduled surgery within 3 weeks of inclusion date, 4) surgery of the upper limbs in the previous three months	MyoSet (MyoPinch, MyoGrip and MyoPlate), BBT, Minnesota Test

O'Neil M.E. et al. ***(2016)***	Observational study	i) to evaluate the inter-instrument reliability and concurrent validity of 3 accelerometer-based motion sensors for measuring PA intensity	Clinical standardized setting	57	CP: hemiplegia (29), diplegia (26), quadriplegia (3)	12.5±3.3	1) GMFCS= I-III, 2) ambulatory children, 3) 6-20 years old, 4) able to follow instructions and protocol directions, 5) able to wear 3 pairs of accelerometers and 1 portable indirect calorimeter.	1) recent musculoskeletal injuries, limiting their PA levels, 2) orthopaedic surgery within the precious 6 months, 3) botulinum toxin or phenol injections within the previous three months, 4) previous unstable medical conditions limiting PA levels, 5) unstable emotional or behavioural status.	NA
Cocker-Bolt P. et al. ***(2016)***	Prospective pre-test/post-test study	To determine the feasibility and use accelerometers before, during and after a CIMT program	Outpatient + Laboratory	12	UCP	4.9 ±1.33	1) UCP, 2) able to use the affected UL as a gross assist during play and self-care activities, 3) no significant developmental delays, 4) ambulatory, 5) no additional health impairments	1) significant intellectual disabilities, 2) seizure disorders, 3) botulinum toxin injections in the previous 6 months	MA2

ADHD: Attention Deficit Hyperactivity Disorder, AHA: Assisting Hand Assessment, BOT-2: Bruininks–Oseretsky Test of Motor Proficiency, CFUS: Caregiver Functional Use Survey, CG: Control Group, CIMT: Constraint-Induced Movement Therapy, DMD: Duchenne Muscular Dystrophia, EDSS: Extended Disability Status Scale, EG: Experimental Group, IQ: Intelligence Quotient, JTHF: Jebsen–Taylor Test of Hand Function, MA2: Melbourne Unilateral Upper Limb Assessment – 2, MAS: Modified Ashworth Score, NP – C: Nieman Pick C, PAFT: Pediatric Arm Function Test, PMAL-R: Pediatric Motor Activity Log – Revised, QUEST: Quality of Upper Extremity Skills Test, UL: upper limb VGT: Video-Game based Therapy.

Forms of Assessment

Index Test

Details about the use of inertial sensors in the two groups of children are reported in Table 3a and 3b, Table 4a and 4b. Application of accelerometers in the selected studies were attributed to several purposes: i) to establish activity intensity cut-points (Kaneko, M., et al. 2016; Phillips, L. R., et al. 2013; Davila E. M. 2011; O'Neil, M. E et al. 2016); ii) to investigate the validity and reliability of specified models of inertial sensors (O'Neil, M. E et al. 2016; Lemmens, R. J. et al. 2015; Phillips, L. R., et al. 2013; Floyd A. G. et al. 2007; Sadeh, A., et al. 1994; Birmingham, A. T., et al. 1985; Strohrmann, C. et al. 2013); iii) to examine the effect of posture, (Birmingham, A. T., et al. 1985) placement and number of the sensors (Strohrmann, C. et al. 2013; Mackintosh, K. A., et al. 2016; Sadeh, A., et al. 1994; Davila E. M. 2011; Bergamini E. et al. 2015; O'Neil, M. E et al. 2016; Lemmens, R. J. et al. 2015); iv) to develop novel monitoring tools, to measure and quantify the symptoms, neurodevelopmental delay and/or the autonomy of patients suffering from chronic disabilities (Coker-Bolt, P. et al. 2017; Gordon, A. M., et al. 2007; Strohrmann, C. et al. 2013; Zoccolillo, L. et al. 2015; Sokal, B., et al. 2015; Bergamini E. et al. 2015; Kaneko, M., et al. 2016; Floyd A. G. et al. 2007; Le Moing A. G., et al. 2016) v) to evaluate the duration and intensity of natural limbs movements, defined motor tasks and tremor (in both groups) (Sokal, B., et al. 2015; Lemmens, R. J. et al. 2015; Zoccolillo, L. et al. 2015; Phillips, L. R., et al. 2013; Gordon, A. M., et al. 2007; Floyd A. G. et al. 2007; Sadeh, A., et al. 1994; Birmingham, A. T., et al. 1985; Bergamini E. et al. 2015; Deutsch, K. M., & Newell, K. M. 2006; MacArthur, B., et al. 2014; Strohrmann, C. et al. 2013; Davila E. M. 2011; O'Neil, M. E et al. 2016; Kaneko, M., et al. 2015; Le Moing A. G., et al. 2016; Graves, L. E., et al. 2008; Dadashi, F., et al. 2016; Mackintosh, K. A., et al. 2016; Coker-Bolt, P. et al. 2017); vi) to assess the efficacy of pre-specified rehabilitation protocols (Zoccolillo, L. et al. 2015; Gordon, A. M., et al. 2007; MacArthur, B., et al. 2014). Compliance with accelerometers were reported in five of the studies (Mackintosh, K. A., et al.

2016; Davila E. M. 2011; O'Neil, M. E et al. 2016; Le Moing A. G., et al. 2016; Bergamini E. et al. 2015) and in general there were well perceived and tolerated. Only in one case (Davila, E. M. 2011) was reported occasional swelling.

Table 3: a) Technical Data for collection phase in Typically Developing Children.

Author	Sensors Number	Sensors Type & Make	Placement	Wearing time	Sample frequency
Birmingham A. T. et al. *(1985)*	2	Accelerometers (Bruel and Kjaer type 4367).	Terminal phalanx of each middle finger	3 min for each hand	5-second epoch
Avi Sadeh et al. *(1994) [Study 1]*	2	Actigraphs (AMA-32, Ambulatory Monitoring, Inc., Ardsley, NY).	Each wrists	2 nights (about 7 hours to night)	1-minute epochs.
Deutsch K. M. et al. *(2006)*	2	Uniaxial wireless accelerometers (Coulbourn T45-10, calibrated on each day of testing).	Dorsal surface of the tip of the distal segment of each index finger.	three 10-s consecutive trials, about 5 s breaks between trials.	200 Hz
Graves L.E.S. et al. (2008)	6	1) Actiheart (Cambridge Neurotechnology Cambridge, UK), 2) 4 uniaxial ActiGraph accelerometers (GT1M, Fort Walton Beach, FL, USA)	1) on the skin at the base of the sternum, 2) on the midaxillary line of the right and left hip and on each forearm proximally from the wrist joint .	60 min	2) 30 Hz
Davila E. M. *(2011)*	2	Actical triaxial AMs (Respironics Co., Inc., Bend, OR, USA).	Dorsal side of each wrist	Full seven days (24 hrs/day).	15-second epoch
Phillips L. R. S. et al. *(2012)*	3 + 1	Triaxial wireless accelerometers GeneActive (Unilever Discover, Colworth, UK) + ActiGraph GT1M (Actigraph, Pensacola, FL, USA).	Each wrist and + right hip (ActiGraph GT1M worn adjacent to the hip mounted GENEA)	Activities: 5 min; Lying supine: 10 min.	GENEA: 80 Hz, ActiGraph GT1M: 1 s epochs.
MacArthur B. et al. *(2014)*	3	Actical accelerometers (Actical, Philips Respironics Co. Inc., Bend, OR).	Each wrist + hip	20 minutes	15-second epoch

Lemmens R. J. M. et al. ***(2015)***	7	Sensor devices, composed by a triaxial accelerometer, triaxial gyroscope, triaxial magnetometer (SHIMMER Research, Dublin, Ireland).	Chest + Dominant and non-dominant arm-hand: on the dorsal side of the hand, of the wrist and on the upper arm	Not specified.	128 Hz
Kaneko M. et al. (2015)	4	wearable sensors composed of three-axis acceleration and three-axis angular velocity sensors (WAA-006, WAA-010, ATR-Promotions, Kyoto, Japan)	both hands and elbows	four motor tasks: 10 s for each task	100 Hz
Dadashi F. et al. (2016)	3	waterproofed IMUs (Physilog III, BioAGM, CH, 3D accelerometer, 3D gyroscope).	2 IMUs placed on the dorsal side and distal end of the forearms, one on the sacrum.	Not specified.	500 Hz
Mackintosh K.A. et al. (2016)	9	triaxial accelerometer (Actigraph wGT3X+, Florida, USA)	on the lateral plane of each ankle, knee, hip, wrist, and centre of the chest.	30 minutes	100 Hz

IMU: Inertial Measurement Unit, Hz: Herz

Table 3: b) Technical Data for analysis phase in Typically Developing Children.

Author	**Accelerometer data comparison**	**Differences between the two hands**	**Data cleaning**	**Threshold** (cutoff frequency of filter applied on raw data)	**Threshold to assess the intensity of arm movement**
Birmingham A. T. et al. ***(1985)***	RMS of tremor amplitude, dominant peak and its frequency.	For rest tremor, amplitude in the dominant hand was significantly lower in adolescence and early adult life than in childhood, for the non-dominant hand the statistically significant difference was sustained to later life. For work tremor, dominant hand frequency declined significantly with age, both hands continue to decline in adulthood.	Frequency analysis of the tremor waveform was filtered to remove frequencies above 50 Hz to prevent alias contamination	50 Hz	NA
Avi Sadeh et al. ***(1994) [Study 1]***	Accelerometric data matched with PSG scoring performed to develop the scoring algorithm: PS probability of sleep	The mean activity level of the dominant wrist was significantly higher than that of the nondominant wrist during PSG-determined sleep (6.84 vs. 6.16), as well as during wakefulness (25.8 vs. 22.3).	NA	NA	NA

Deutsch K. M. et al. *(2006)*	The peak frequency within two frequency bands (5-15 Hz and 15-30 Hz) and the proportion of power exhibited at the peak frequency determined (based on power spectral density calculated using Welch's averaged periodogram method).	The peak frequency of the finger of the dominant hand (21.4 Hz) was higher than nondominant hand (20.7 Hz) in the 15-30 Hz frequency band. No significant differences in proportion of power exhibited at peak frequency within the 5-15 Hz of postural tremor as a function of age, hand dominance or hand configuration. Postural tremor of nondominant hand was significantly more regular than dominant hand.	Band-pass filtered	1 Hz - 50 Hz	NA
Graves L.E.S. et al. (2008)	Means and standard deviations of activity counts (counts/min)	Activity of the dominant limb was significantly greater than non-dominant during tennis and bowling ($P < 0.001$) and non-dominant limb activity was significantly greater during boxing than bowling or tennis ($P < 0.001$). Activity counts from the left wrist for tennis and boxing ($r = 0.710$ and 0.744, $P \cdot 0.01$) and the right wrist for boxing ($r = 0.586$, $P < 0.05$) were significantly correlated with EE.	Band pass filtering	0.21–2.28 Hz	NA

Davila E. M. ***(2011)***	Data Trasformation: AEE, Time. Data Summarization Characteristics: Bouts Duration, Intensity Thresholds.	No statistical differences between outcome variables for any bout duration (1, 5, 10 minutes) within L and MV intensity categories between AMs (D versus ND, LW versus RW) or model (1R versus 2R). Dominant and RW AMs were no-significantly higher than ND and LW, respectively, within MVPA intensity. In contrast, ND and LW AMs were non-significantly higher than D and RW within L intensity PA. Identical results within gender.	Quantity control checks were performed to identify periods on non-wear.	NA	Light (AEE < 0.05 kcals/kg/min), moderate (0.05 < AEE < 0.09 kcals/kg/min), vigorous (AEE ≥ 0.10 kcals/kg/min).
Phillips L. R. S. et al. ***(2012)***	VM with gravity-substracted.	Both sides demonstrated good criterion validity (right: r=0.9, left: r=0.91) and good concurrent validity (right: r=0.83, left: r=0.845). ROC analysis proved GENEA monitors able to successfully discriminate among all intensity levels.	NA	NA	Sedentary (< 1.5 METs), light (1.5-2.99 METs), moderate (3-5.99 METs) and vigorous (≥ 6 METs). The accelerometer counts for activities were coded into binary indicator variables (0 or 1) based on intensity.
MacArthur B. et al. ***(2014)***	Percentage of time in MVPA calculated by summing the number of 15-second intervals in which the activity counts were ≥ 574 counts/15 seconds.	The accelerometers placed on the wrists did not find differences in the conditions in percentage MVPA (right: 48.8±29.5%, left: 47.6±28.8%).	NA	NA	MVPA: activity counts ≥ 574 counts/15 seconds.

Lemmens R. J. M. et al. *(2015)*	ICC parameter (based on VM).	Within-subject reliability calculated for the 2 arm hands separately, median ICCs ranged between 0.68-0.92. Between subject reliability for the 2 arm hands separately, median ICCs ranged between 0.61-0.90.	Zero time-phase, low-pass filtered	1.28 Hz	NA
Kaneko M. et al. (2015)	Postural stability of the hands and elbows, rotational speed, mirror movement, two parameters of bimanual symmetry, compliance	All indices had a tendency to increase with age.	Low-pass filter	6 Hz	NA
Dadashi F. et al. (2016)	Average propulsive phases of right and left arms, pull and push phases (Δpull, Δpush), sum of aerial recovery and entry catch phases (ΔNProp), index of coordination (IdC).	By increasing the velocity, the duration of arm under-water phases (Δpull + Δpush) and accordingly IdC did not change significantly. G2 group used 2,8% lower catch-up pattern ($P<0,01$) by increasing the arm under-water phases ($P<0.016$) and using 6.5 more arm stroke ($P<0.001$). No changes in the stroke length and cycle velocity variation were observed ($P > 0.22$).	NA	NA	NA
Mackintosh K.A. et al. (2016)	Mean and variance of the accelerometer counts in each 15 s window. These extracted features were used as inputs into the ANNs, a specific type of machine learning model. RMSE.	The ANNs for left and right wrist accelerometers had a lower correlations with predicted EE. No significant differences in RMSE analysis. Despite significant advantages in terms of compliance, they could lead to potentially marginal losses in EE prediction accuracy.	NA	NA	1,4% of collected data were removed when EE < 0,5 MET (measured with MetaMax 3B)

AEE: Activity Energy Expenditure, AM: Activity Monitors, ICC: Intraclass Correlation Coefficient, IMU: Inertial Measurement Unit, MET: Molecular Electronic Transducers, PSG: Polysomnography, RMS: Root Mean Square, RMSE: Root Mean Square Error, ROC: Receiver Operating Characteristic, VM: Vector Magnitudes.

Table 4: a) Technical Data for collection phase in Children with Neurodevelopmental Disorders.

Author	Sensors Number	Sensors Type & Make	Placement	Wearing Time	Sample frequency
Floyd A. G. et al. *(2007)*	2	Piezoresistive uniaxial accelerometers with linear sensitivities of 4.5 mV/g in the biological tremor range (0-25 Hz)	Over the dorsum of both hands	Multiple recording and total recording time lasted 1-2 h	300 Hz
Gordon A. M. et al. *(2007)*	2	Accelerometers (Manufacturing Technology Inc. Fort Walton Beach, FL, model 7164)	Each wrist	During the AHA test session	10 Hz
Strohrmann C. et al. *(2013)*	10	ETH Orientation Sensor (ETHOS)= IMU composed by a 3D accelerometer, a 3D gyroscope and a 3D digital compass. Not commercially available.	Upper (wrists and upper arms) and lower extremities (upper legs and feet) and the trunk.	1h, once per week over a course of four weeks.	100 Hz
Zoccolillo L. et al. *(2015)*	5	Wireless triaxial accelerometers (Trigno, Delsys®).	Posterior part of forearms, of shanks and of lower trunk in correspondence of the centre of mass (L2-L3).	During 5 continuous minutes of video-game based therapy and 5 minutes of CT.	Not specified
Sokal B. et al. *(2015)*	2	Biaxial wireless accelerometers (Model 71256, Actigraph, Pensacola, FL)	Dorsal side of both wrists just above the styloid process	During waking hours for at least 9 hours daily for 3 consecutive days after the testing session.	10 Hz, integrated over a user-specified epoch (2 s).
Bergamini E. *(2014)*	3	IMUs (Opal, APDM Inc., Portland, Oregon, USA).	Both wrists and backrest of the wheelchair.	Time was manually recorded. Total time not reported.	128 Hz

Kaneko M. et al. (2016)	4	Acceleration and angular velocity sensors (WAA-006, WAA-010, ATR-Promotions, Kyoto, Japan)	Both hands and elbows	two motor tasks (imitative motor task and a maximal-effort motor task): 10 s for each task	100 Hz
Le Moing A.-G. et al. (2016)	2	Watch-like devices contained a three-axis accelerometer, a three-axis gyroscope, and a three-axis magnetometer	On each wrist	at least 30 minutes to complete all the tasks, without concerning potential resting period	NA
O'Neill M.E. et al. (2016)	6	1) StepWatch activity monitor (uniaxial), 2) Actigraph GT3X (triaxial), 3) BodyMedia SenseWear Pro Armband (triaxial).	1) superior to the left/right malleolus, 2) on a waist elastic belt superior to the right/left iliac crest, 3) dorsal side of each upper arm at the midbelly of the triceps muscle	During each data collection, lasting 2-2,5 hours	1 s for ActiGraph, 3 s for StepWatch, and 60 s for SenseWear.
Coker-Bolt P. et al. (2017)	2	Triaxial Actigraph GT9X Link (Actigraph, Pensacola, FL)	on each wrist	6 hours a day before and after the CIMT program (tot: 12 hours)	30 Hz

Table 4: b) Technical Data for analysis phase in Children with Neurodevelopmental Disorders.

Author	**Accelerometer data comparison**	**Differences between the two hands**	**Data cleaning**	**Threshold** (cutoff frequency of filter applied on raw data)	**Threshold to assess the intensity of arm movement**
Floyd A. G. et al. *(2007)*	Side-to-side relationship of tremor amplitude, peak tremor frequencies and amplitude variability.	Action tremor amplitudes were relatively symmetric between the dominant and non-dominant hands, postural tremor was not symmetric bilaterally (3 of 8 patients were unilateral), amplitudes of bilateral cases correlated within subjects.	In the FTN trials only, frequencies below 2 Hz were excluded	2 Hz	NA
Gordon A. M. et al. *(2007)*	Percentage of hand use (activity counts)	The percentage of use of involved extremity remain the same in controls, 70% of the task performance, while increased from 62.6 to 77.8% for the children who received HABIT (not correlate with the change in AHA scores). Use of the non-involved extremity remained the same across testing sessions in both groups.	NA	NA	NA

Strohrmann C. et al. ***(2013)***	TIME, mean value of MI, MIV, DF, SM, ARE, RANG, ArmSync, gait parameters (all based on VM).	MIV is larger for the unaffected hand, the energy associated to the dominant frequency of the affected hand vs. unaffected hand was much lower, the SM parameter of the unaffected side vs. affected side was twofold.	Low-pass filtered	45 Hz	NA
Zoccolillo L. et al. ***(2015)***	RMS of acceleration.	Hemiparetic side was moved less than healthy side. In VGT the paretic side was moved -20±13% less than the other side, while this difference was not significant in CT (-10±28%).	Low-pass filtered and after the mean substraction for removing the contribution of gravity acceleration.	20 Hz	NA
Sokal B. et al. ***(2015)***	Duration SV, duration ratio SV, intensity SV, intensity ratio SV.	Partecipants moved their more-affected arm for 55.7% and their less-affected arm for 64.9%, ratio 0.86. The intensity of more-affected arm was 41.3 counts/s and for less-affected arm was 60.5, ratio 0.71.	Segments when partecipants appeared to have removed the accelerometers were removed.	NA	Raw values for each 2 s recording epoch were dichotomized around a low threshold (i.e., 2) with above-threshold values set to a positive costant and at- or below-threshold values set to zero.
Bergamini E. ***(2014)***	Symmetry index, a peak of the acceleration magnitude and CV (all based on VM).	Symmetry index: - CG: ES2 (48.92%) and ES3 (47.86%), - EG: ES2 (47.77%) and ES3 (48.62%). These values indicate good symmetry.	Low-pass filtered	12 Hz	NA

Kaneko M. et al. (2016)	Rotational speed, mirror movement, postural stability of rotating elbow, temporal change of rotational size in each index, bimanual symmetry, compliance.	All scores of ADHD children was lower than TD children. In bimanual symmetry the score of ADHD children increased with age and was significantly different to TD aged 8 and 10 years old. The variability of children's score in compliance and temporal change of rotational size in ADHD vs. TD was larger.	low-pass filter	6 Hz	NA
Le Moing A.-G. et al. (2016)	Norm of the angular velocity, ratio of the vertical component of the acceleration, model-based computed power, elevation rate	Not find any side effect between the dominant and non-dominant hands. Patients performed better with their dominant side but this was not statistically significant, due to the small size of the population and the advanced stage of the disease.	NA	NA	NA
O'Neill M.E. et al. (2016)	Median (IQR) evaluated and compared between right and left side for each parameter and each device, ICC, CIs	Each accelerometer is stable in data collection on both sides, indicating that movement asymmetries may not influence PA measures. Because all 3 accelerometer models exhibited excellent inter-instrument reliability for measuring PA in a variety of real-world activities in TD, it may be appropriate also for CP to wear accelerometers on the right side.	NA	NA	NA

Coker-Bolt P. et al. (2017)	Active duration, mean activity count, use ratio and magnitude ratio (all based on VM, down-sampled to 1 Hz).	Significant increase in the duration and mean activity count of affected upper limb use during each camp day and in three of five days in comparison to pre-test data, respectively. No significant changes in all scores pre- vs. post-CIMT.	NA	NA	Upper limb activity when the vector sum activity count > 0.

ARE: Average Rotation Energy, ArmSync: Synchrony of Arm Movement, CIs: Confidence Intervals, CT: Conventional Therapy, CV: Intercycle Variability, DF: Dominant Frequency, FTN: finger – to – nose, IQR: Interquartile Range, MI: Movement Intensity, MIV: Movement Intensity Variation

Accelerometer Type, Site and Duration

Included studies were grouped by: i) the type of the sensors, ii) the placement on the body, depending on whether they were placed only on the ULs or on other parts of the body and iii) the wearing time. When reported, several different inertial sensors were variably used: tri-axial (n. 6 studies (Phillips, L. R., et al. 2013; Davila E.M. 2011; Zoccolillo, L. et al. 2015; O'Neil, M. E et al. 2016; Mackintosh, K. A., et al. 2016; Coker-Bolt, P. et al. 2017), two-axial (n. 1 study, Sokal, B., et al. 2015), and uni-axial accelerometers (n. 3 studies Floyd A.G. et al. 2007; Deutsch, K. M., & Newell, K. M. 2006; Graves, L. E., et al. 2008). Some studies used IMUs (n. 6, Bergamini E. et al. 2015; Le Moing A. G., et al. 2016; Dadashi, F., et al. 2016; Kaneko M. et al. 2015; Kaneko, M., et al. 2016; Lemmens, R. J. et al. 2015), or ETHOS (Strohrmann, C. et al. 2013), namely an IMU composed by a 3D accelerometer, a 3D gyroscope and a 3D digital compass.

The placement of the sensors in the UL was mainly in the dorsal side of both wrists (Sokal, B., et al. 2015; Gordon, A. M., et al. 2007; Floyd A.G. et al. 2007; Sadeh, A., et al. 1994; Davila E. M. 2011; Le Moing A. et al. (2016); Coker-Bolt, P. et al. 2017), while in few cases in terminal phalanx or on the tip of the fingers (Birmingham A.T. et al. 1985; Deutsch, K. M., & Newell, K. M. 2006) or on each elbow (Kaneko M. et al. 2015; Kaneko, M., et al. 2016). Additional sensors were placed on the hip (Phillips, L. R., et al. 2013; MacArthur, B., et al. 2014), on the back (Dadashi, F., et al. 2016), on lower extremities (Zoccolillo, L. et al. 2015) and on the trunk (Strohrmann, C. et al. 2013; Mackintosh, K. A., et al. 2016) and on the chest (Lemmens, R. J. et al. 2015) or externally on the backrest of a wheelchair (Bergamini E. et al. 2015). The wearing time varied across studies (Table 3a and 4a) from a daily duration of 30 minutes to 9 hours, a from 1 to 7 days. The type of activities performed were various including rest position, walking, writing, assessment or intervention.

Accelerometer Data Collection and Analyses

Data were collected either as the summed acceleration counts over a specified time, or as dichotomized data representing the duration of active and inactive periods (Sokal, B., et al. 2015; Coker-Bolt, P. et al. 2017). Various approaches were reported for defining a unit of arm activity, with data capture epochs varying from one second to one minute. For the analysis, five studies use a 3-D resultant vector (usually known as "vector magnitude") calculated by applying Pythagoras' rule to each time point of the x-, y- and z-components of the signals of each sensor (Bergamini E. et al. 2015; Strohrmann, C. et al. 2013; Lemmens, R. J. et al. 2015; Phillips, L. R., et al. 2013; Coker-Bolt, P. et al. 2017), three studies use the tremor amplitude obtained by accelerometer data (Floyd A. G. et al. 2007; Birmingham A. T. et al. 1985; Deutsch, K. M., & Newell, K. M. 2006), and another one use the root mean square (RMS) of acceleration computed on the signals for each device (Zoccolillo, L. et al. 2015).

In two studies the participants were asked to keep a diary to record wear time and modes and bouts of activity to assist in analysis of accelerometer data (Sokal, B., et al. 2015; Davila E. M. 2011).

The included studies could be grouped by outcome measures for accelerometer data comparison between ULs. Four studies compared UL movement using the intensity of arm movement. Sokal, B., et al. 2015 evaluated the asymmetry as the ratio of the intensity of more-impaired to less-impaired arm movement, whereas the intensity of arm movement was quantified by dividing the sum of the raw recordings by the sum of the threshold-filtered recordings for each arm. Three studies compared the intensity of arm movement after defining the intensity categories from light (L) to moderate-to-vigorous (MV) PA (Sedentary (< 1.5 METs), light (1.5-2.99 METs), moderate (3-5.99 METs) and vigorous (≥ 6 METs)). (Phillips, L. R., et al. 2013; Davila E. M. 2011; MacArthur, B., et al. 2014).

The comparison of ULs movement amount was variably done by using Root Mean Square (RMS) (Zoccolillo, L. et al. 2015), by means and standard deviations of activity counts (counts/min) (Graves, L. E., et al. 2008; Mackintosh, K. A., et al. 2016; Coker-Bolt, P. et al. 2017), by the activity counts along with a synchronized video to determine the percentage of time each hand was used while performing the activities (Gordon, A. M., et al. 2007) or by elaborating an algorithm matching the actigraph data with polysomnography (PSG) (Sadeh, A., et al. 1994).

Other studies evaluated the asymmetry of arm movement on the basis of features extracted from sensor data by considering the differences in peak acceleration measured by the IMUs located on the wrists (Bergamini E. et al. 2015), or the correlation coefficient of acceleration along the Z axis between the right hand and the left hand, and time delay of acceleration in the Z axis between the right hand and left hand (Kaneko M. et al. 2015, Kaneko, M., et al. 2016), among different sensors put on the feet, the upper legs and the wrists (Strohrmann, C. et al. 2013), or creating model-based computation of the scalar product of the torque and the angular velocity (W/kg) of the wrist (Le Moing A.G. et al. 2016).

Type of Outcome Measure and Reference Standard

Studies included both subjective and objective measures of the physical activity and energy expenditure. Two of the studies did not use a reference standard (Deutsch, K. M., & Newell, K. M. 2006; Dadashi, F., et al. 2016) and six used Demographic Questionnaires (MacArthur, B., et al. 2014; Lemmens, R. J. et al. 2015; Birmingham, A. T., et al. 1985; Davila, E. M. 2011; Kaneko, M., et al. 2015; Kaneko, M., et al. 2016). Six of the remaining thirteen studies used a reference standard, which was unsuitable to correctly classify the bimanual asymmetry: portable gas analyser Cosmed K4b2 (O'Neil, M. E et al. 2016; Phillips, L. R., et al. 2013), portable indirect calorimeter MetaMax 3B (Graves, L. E., et al. 2008; Mackintosh, K. A., et al. 2016) and heart rate recordings (Actiheart) (Graves, L. E., et al. 2008); Arm Crank Ergometer

including LODE (Bergamini, E. et al. 2015), polysomnography (PLMS) and electroencephalogram (EEG) (Sadeh, A., et al. 1994). Other unsuitable reference standard were different types of questionnaires, such as the Expanded Disability Status Scale (EDSS) (Floyd, A. G. et al. 2007); Children's Hand-use Experience Questionnaire (CHEQ) (Coker-Bolt, P. et al. 2017), Childhood Autism Rating Scale (CARS) and Physical Activity Enjoyment Scale and Pediatric Quality of Life Inventory 4.0 (PedsQL 4.0) (MacArthur, B., et al. 2014), Pediatric Evaluation Disability Inventory (PEDI) (Coker-Bolt, P. et al. 2017) and questionnaires about the physical activity behaviour, investigating about the medical status of health, and rehabilitation services (O'Neil, M. E et al. 2016) and the gaming experience (Graves, L. E., et al. 2008).

In seven of the included studies, children's motor function was extensively and reliably assessed by experts, using direct observations of uni- and bimanual tasks, such as Quality of Upper Extremity Skill Test (QUEST) (Zoccolillo, L. et al. 2015; DeMatteo, C. et al. 1993), Assisting Hand Assessment (AHA) (Gordon, A. M., et al. 2007; Krumlinde-Sundholm, L., Holmefur, M., Kottorp, A., & Eliasson, A. C. 2007), Melbourne Unilateral Upper Limb Assessment-2 (MA2) (Coker-Bolt, P. et al. 2017; Randall, M., Johnson, L., & Reddihough, D. 1999). Other possible appropriate reference standards were Bruininks Oseretsky Test of Motor Proficiency (BOT-2), Caregiver Functional Use Survey (CFUS) and the Jebsen Taylor Test of Hand Function (JTHF) (Gordon, A. M., et al. 2007; Deitz, J. C., Kartin, D., & Kopp, K. 2007; Jebsen, R. H. 1969), Pediatric Motor Activity Log – Revised (PMAL-R) (Sokal, B., et al. 2015), ABILHAND-Kids (Zoccolillo, L. et al. 2015; Arnould, C., Penta, M., Renders, A., & Thonnard, J. L. 2004), Box and Block Test (BBT), Minnesota Test, and MyoSet (Le Moing A. G., et al. 2016). In other studies skilled professionals analysed video recordings, as using Pediatric Arm Function Test (PAFT) (Sokal, B., et al. 2015; Uswatte, G. et al. 2012), and Motor Capacity Assessment (Strohrmann, C. et al. 2013), followed by the estimation of a final general score. Motor Capacity Assessment entailed 10 selected predefined motor tasks were from established and validated motor assessments, namely

JTHF, the Graded and Redefined Assessment of Strength, Sensibility and Prehension (GRASSP), Nine-Hole Peg Test (NHPT), and Timed Up and Go. All are standardized tests, aiming to assess the function of the two ULs separately. For this reason, these tests and their scores can be used as a comparison measure, to evaluate and to interpret the results of the analysis of accelerometer data. These tests are currently used in the clinical practice, therefore they are characterized by standardized score ranges, which correspond with different motor capabilities and dexterity levels. As a consequence, the reported results for each participant can be reliably used as a reference, to describe intensity activity accelerometers brackets and motor patterns, recorded by the movement sensors, allowing to do inferences about their reliability.

It was shown that the motor capacity measure assessed by the PAFT was significantly correlated with the intensity and duration of more-affected arm measured in laboratory (Sokal, B., et al. 2015). Other interesting relationships were assessed by QUEST and AHA. The significant changes in the QUEST after a videogame based therapy (VGT) were related to higher quantitative movements during the VGT (Zoccolillo, L. et al. 2015) and also the changes detected by AHA after an hand–arm bimanual intensive therapy (AHA) were detected by accelerometers used during AHA assessments (Gordon, A. M., et al. 2007). Another study has shown an interesting significant correlation between data acquired by IMUs and motor capacity assessment, a rating assessment performed by independent expert raters on video recordings (Strohrmann, C. et al. 2013). In addition, accelerometer data have shown a high significant correlation with "MyoSet", a tool for measuring patients' ability to move their hand using finger and wrist flexors and extensor by using the 'MoviPlate' assessment, which consists in hitting alternatively two targets of different heights, placed in the sagittal plane for 30 seconds and Block and Box Test (Le Moing A. G., et al. 2016). Finally, the data collected by accelerometers before, during and after a 1-2 weeks long CIMT-program were consistent with the clinical scores (MA2), in terms of potential improvements

or decreases in the quality of affected UL movements (Coker-Bolt, P. et al. 2017).

Use of the inertial sensors in TD children

The most important and common conclusion that can be derived from the studies of inertial sensors in TD children is that there were minimal differences observed for any variable between monitors, worn on both ULs, during gross motor activities. For example, there were no differences in detecting sleep period collecting data either from the dominant or non dominant wrists during sleep-wake cycle (Sadeh, A., et al. 1994). The measurement of the upper body movement, between limbs, was similar in the different activities such as active videogames sessions and outdoor play in terms of differentiating moderate- to vigorous physical activity (MacArthur, B., et al. 2014).

In a study of 44 TD participants performed a series of activities representative of daily life, such as active videogames, lying supine, seated DVD viewing and walking and running at various speeds (Phillips, L. R., et al. 2013). All were symmetrical activities, concerning the upper body segment movements and this was reflected in no significant differences between the accelerometer output for the wrist worn Actigraphs across age groups.

Two studies of wrist worn sensors in TDC confirmed the minimal variability in data between wrist sensors by side which were not impacted by gender, intensity of the activity (low and high), mode of the exercise (free living condition or pre-specified physical activities) (Davila, E. M. 2011). This was confirmed in a second study were repeated tasks analysed using repeated measures analysis of covariance, controlling for differences in condition and condition order confirmed no differences between sides (MacArthur, B., et al. 2014). It can therefore be concluded that in TD the wrist actigraphy, providing reliable recording of each wrist, could become a promising tool for measurement of bimanual activities and for assessing energy expenditure and physical activity (Graves, L. E., et al. 2008).

On the contrary, when asymmetric UL bimanual tasks (such as drinking from a cup", "eating with knife and fork", "combing hair" and "opening a zipper") were considered appropriate differences between the two ULs were found (Lemmens, R. J. et al. 2015) with a greater variability in movement trajectories in complex task compared to simple ones. Consequently, accelerometers showed a good reliability also in the studies which explored the physiological differences between the two UL. For instance, the discrepancies in the trajectories and intensity of arms movements had been evaluated with inertial sensors in two groups of swimmers, to define the variability patterns of their techniques. Various different movement descriptors made possible to identify and compare groups of different performing levels. As a consequence, inertial sensors could be used also as a tool to refine motor skills (Dadashi, F., et al. 2016). Moreover, when the features of physiologic movements throughout different age groups were detected, it was possible to identify accurately the age-related changes occurring in the physiologic tremor frequency profiles (Birmingham, A. T., et al. 1985; Deutsch, K. M., & Newell, K. M. 2006) and in soft neurological signs (SNS) (Kaneko, M., et al. 2015). The soft neurological signs are minor neurological findings, which are likely to appear in the motor performance of typically developing young children, disappearing as the child grows up.

Use of inertials sensors in children with NDDs

From the analysis of this systematic review, it can be summarized that the inertial sensors are able to differentiate the different trends between TD and NDDs children. This observation can be drawn especially due to two studies, in which the same evaluation of the hand pronation and supination is administered both in a group of TD children and in a sample of children with a diagnosis of ADHD. The obtained movement development curves for the two groups of children were compared. Indices such as bimanual symmetry, rotational speed and postural stability of the both hands were lower for children

with NDDs than for the TD children group (Kaneko, M., et al. 2015; Kaneko, M., et al. 2016).

Similarly, studies of children with NDDs identified a greater asymmetry in the usage of the UL, in contrast to the typical bimanual cooperation, characterized by performing activities using the two arms in a balanced fashion. Studies concerning children with UCP assessed the use of accelerometry both in uni- and bimanual tasks to evaluate the asymmetry between the two limbs (Coker-Bolt, P. et al. 2017; Zoccolillo, L. et al. 2015; Gordon, A. M., et al. 2007; Strohrmann, C. et al. 2013; Sokal, B., et al. 201515). Significant differences between the impaired and the unimpaired UL accelerometer data arose from all of the included articles. It is often observed that the natural tendency would be to over-compensate with a higher usage of their non-involved extremity, as much as higher the severity level of the impairment. The analysis of the collected accelerometer data matched this asymmetric trend and therefore reliability was confirmed.

In one study, in order to measure the synchrony of the two arms, was added wheel chair driving, in addition to walking and stair climbing. A smoother pattern was detected by accelerometry in the unimpaired hand, compared to the other side, meaning that a performance characterized by several different frequencies indicates a more uncontrolled movement (Strohrmann, C. et al. 2013). The changes in clinical outcome measures such as QUEST, AHA, MA2 and PAFT after treatment, VGT, HABIT and CIMT, respectively, were related to higher intensities of activity and frequency of use recorded by actigraphs (Zoccolillo, L. et al. 2015; Gordon, A. M., et al. 2007; Coker-Bolt, P. et al. 2017; Sokal, B., et al. 2015).

It can therefore be suggested that accelerometer data analysis is able to measure and describe the differences in the severity of a wide range of pathological conditions. For instance, in movement disorders such as tremor, dystonia, chorea and myoclonus secondary to neurodegenerative disorders (e.g. Nyeman Pick-C) the accelerometry is able to quantify the rate of them with data related to EMG data (Floyd, A. G. et al. 2007). These data may provide important

methods to determine the severity and progression of pathology and the outcome of treatments (Floyd, A. G. et al. 2007).

The findings of these studies also suggested that it should be considered that the necessary amount of quantity and quality of the movement may be related to the complexity of the task itself, since the laboratory approach did not necessarily include the tasks, daily performed in a dwelling environment (Strohrmann, C. et al. 2013). This point is further strengthened in one study in which the main detected limitation was the clinical setting, even if the protocol was designed to resemble real-world activities as much as possible (O'Neil, M. E et al. 2016).

In addition, accelerometer data analysis can provide interesting insights in the study of technique patterns and overall test performance, as shown for TD children (Dadashi, F., et al. 2016). For instance, indices such as the bilateral symmetry index can help in the study of biomechanical characterization of the wheelchair propulsion, identifying potential strengths and weaknesses (Bergamini, E. et al. 2015).

A controversial aspect is the relationship between the accelerometer data and the daily use of the upper limbs in hemiplegia. Several studies in adults with hemiplegia due to chronic stroke report strong correlations between amount of movement and use of the more-affected arm (24–27) while in one study in children (Sokal, B., et al. 2015) found no correlation between the amount of movement of the more-affected arm in daily life with its amount to use, suggesting that children differ from adults in this respect. Further studies are needed to explore this topic, and in particular observational studies comparing the use of the two hands in children with hemiplegia with those of TD children. An important aspect could be the presence of mirror movements in hemiplegic children that can reduce the difference in the amount of use between the more affected hand and the less affected hand. Another important aspect to be further investigated is the use of accelerometers in detecting changes after experimental trainings in upper limb is the monitoring of training. The results in this field are very promising (Gordon, A. M., et al. 2007; Coker-Bolt, P. et

al. 2017) but some results are controversial suggesting that the changes detected by the actigraphs may not be related to those by the clinical scale (Gordon, A. M., et al. 2007).

4.2.3 References of the review:

- Arnould, C., Penta, M., Renders, A., & Thonnard, J. L. (2004). ABILHAND-Kids A measure of manual ability in children with cerebral palsy. *Neurology*, *63*(6), 1045-1052.
- Beaton, A. V. (2003). The nature and determinants of handedness.
- Bergamini, E., Morelli, F., Marchetti, F., Vannozzi, G., Polidori, L., Paradisi, F., ... & Delussu, A. S. (2015). Wheelchair propulsion biomechanics in junior basketball players: A method for the evaluation of the efficacy of a specific training program. *BioMed research international*, *2015*.
- Birmingham, A. T., Wharrad, H. J., & Williams, E. J. (1985). The variation of finger tremor with age in man. *Journal of Neurology, Neurosurgery & Psychiatry*, *48*(8), 788-798.
- Cioni, G., & Sgandurra, G. (2013). Normal psychomotor development. *Handbook of clinical neurology*, *111*, 3-15.
- Coker-Bolt, P., Downey, R. J., Connolly, J., Hoover, R., Shelton, D., & Seo, N. J. (2017). Exploring the feasibility and use of acceleromters before, during, and after a camp-based CIMT program for children with cerebral palsy. *Journal of Pediatric Rehabilitation Medicine*, *10*(1), 27-36.
- Dadashi, F., Millet, G. P., & Aminian, K. (2016). Front-crawl stroke descriptors variability assessment for skill characterisation. *Journal of sports sciences*, *34*(15), 1405-1412.
- Davila, E. M. (2011). *A comparison of bilaterally wrist-worn accelerometers on measures of free-living physical activity in adolescents* (Doctoral dissertation, Montana State University-Bozeman, College of Education, Health & Human Development).
- Deitz, J. C., Kartin, D., & Kopp, K. (2007). Review of the Bruininks-Oseretsky test of motor proficiency, (BOT-2). *Physical & occupational therapy in pediatrics*, *27*(4), 87-102.
- DeMatteo, C., Law, M., Russell, D., Pollock, N., Rosenbaum, P., & Walter, S. (1993). The reliability and validity of the Quality of Upper Extremity Skills Test. *Physical & Occupational Therapy in Pediatrics*, *13*(2), 1-18.
- Deutsch, K. M., & Newell, K. M. (2006). Age-related changes in the frequency profile of children's finger tremor. *Neuroscience letters*, *404*(1), 191-195.
- Duruöz, M. T. (Ed.). (2014). *Hand function: a practical guide to assessment*. Springer Science & Business Media.
- Floyd, A. G., Yu, Q. P., Piboolnurak, P., Wraith, E., Patterson, M. C., & Pullman, S. L. (2007). Kinematic analysis of motor dysfunction in Niemann-Pick type C. *Clinical neurophysiology*, *118*(5), 1010-1018.
- Gordon, A. M., Schneider, J. A., Chinnan, A., & Charles, J. R. (2007). Efficacy of a hand–arm bimanual intensive therapy (HABIT) in children with hemiplegic cerebral palsy: a randomized control trial. *Developmental Medicine & Child Neurology*, *49*(11), 830-838.

- Graves, L. E., Ridgers, N. D., & Stratton, G. (2008). The contribution of upper limb and total body movement to adolescents' energy expenditure whilst playing Nintendo Wii. *European journal of applied physiology*, *104*(4), 617.
- Jebsen, R. H. (1969). An objective and standardized test of hand function. *Arch. Phys. Med. Rehabil.*, *50*, 311-319.
- Kaneko, M., Yamashita, Y., Inomoto, O., & Iramina, K. (2015). Soft neurological signs in childhood by measurement of arm movements using acceleration and angular velocity sensors. *Sensors*, *15*(10), 25793-25808.
- Kaneko, M., Yamashita, Y., & Iramina, K. (2016). Quantitative Evaluation System of Soft Neurological Signs for Children with Attention Deficit Hyperactivity Disorder. *Sensors*, *16*(1), 116.
- Krumlinde-Sundholm, L., Holmefur, M., Kottorp, A., & Eliasson, A. C. (2007). The Assisting Hand Assessment: current evidence of validity, reliability, and responsiveness to change. *Developmental Medicine & Child Neurology*, *49*(4), 259-264.
- Le Moing, A. G., Seferian, A. M., Moraux, A., Annoussamy, M., Dorveaux, E., Gasnier, E., ... & Servais, L. (2016). A Movement Monitor Based on Magneto-Inertial Sensors for Non-Ambulant Patients with Duchenne Muscular Dystrophy: A Pilot Study in Controlled Environment. *PloS one*, *11*(6), e0156696.
- Lemmens, R. J., Seelen, H. A., Timmermans, A. A., Schnackers, M. L., Eerden, A., Smeets, R. J., & Janssen-Potten, Y. J. (2015). To What Extent Can Arm–Hand Skill Performance—of Both Healthy Adults and Children—Be Recorded Reliably Using Multiple Bodily Worn Sensor Devices?. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, *23*(4), 581-590.
- MacArthur, B., Coe, D., Sweet, A., & Raynor, H. (2014). Active videogaming compared to unstructured, outdoor play in young children: percent time in moderate-to vigorous-intensity physical activity and estimated energy expenditure. *Games for health journal*, *3*(6), 388-394.
- Mackintosh, K. A., Montoye, A. H. K., Pfeiffer, K. A., & McNarry, M. A. (2016). Investigating optimal accelerometer placement for energy expenditure prediction in children using a machine learning approach. *Physiological measurement*, *37*(10), 1728.
- Meyer, M. R. U., Baller, S. L., Mitchell, S. M., & Trost, S. G. (2013). Comparison of 3 accelerometer data reduction approaches, step counts, and 2 self-report measures for estimating physical activity in free-living adults. *Journal of Physical Activity and Health*, *10*(7), 1068-1074.
- Mukhopadhyay, S. C. (2015). Wearable sensors for human activity monitoring: A review. *IEEE sensors journal*, *15*(3), 1321-1330.
- Noorkõiv, M., Rodgers, H., & Price, C. I. (2014). Accelerometer measurement of upper extremity movement after stroke: a systematic review of clinical studies. *Journal of neuroengineering and rehabilitation*, *11*(1), 144.

- O'Neil, M. E., Fragala-Pinkham, M., Lennon, N., George, A., Forman, J., & Trost, S. G. (2016). Reliability and validity of objective measures of physical activity in youth with cerebral palsy who are ambulatory. *Physical therapy*, *96*(1), 37-45.
- Oftedal, S., Bell, K. L., Davies, P. S., Ware, R. S., & Boyd, R. N. (2014). Validation of accelerometer cut points in toddlers with and without cerebral palsy. *Medicine & Science in Sports & Exercise*, *46*(9), 1808-1815.
- Oldfield, R. C. (1971). The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia*, *9*(1), 97-113.
- Phillips, L. R., Parfitt, G., & Rowlands, A. V. (2013). Calibration of the GENEA accelerometer for assessment of physical activity intensity in children. *Journal of science and medicine in sport*, *16*(2), 124-128.
- Randall, M., Johnson, L., & Reddihough, D. S. (1999). *The Melbourne assessment of unilateral upper limb function*. Occupational Therapy Department, Royal Children's Hospital, Melbourne.
- Sadeh, A., Sharkey, M., & Carskadon, M. A. (1994). Activity-based sleep-wake identification: an empirical test of methodological issues. *Sleep*, *17*(3), 201-207.
- Sokal, B., Uswatte, G., Vogtle, L., Byrom, E., & Barman, J. (2015). Everyday movement and use of the arms: Relationship in children with hemiparesis differs from adults. *Journal of pediatric rehabilitation medicine*, *8*(3), 197-206.
- Strohrmann, C., Labruyère, R., Gerber, C. N., van Hedel, H. J., Arnrich, B., & Tröster, G. (2013). Monitoring motor capacity changes of children during rehabilitation using body-worn sensors. *Journal of neuroengineering and rehabilitation*, *10*(1), 83.
- Steenhuis, R. E., Bryden, M. P., Schwartz, M., & Lawson, S. (1990). Reliability of hand preference items and factors. *Journal of Clinical and Experimental Neuropsychology*, *12*(6), 921-930.
- Trost, S. G., Zheng, Y., & Wong, W. K. (2014). Machine learning for activity recognition: hip versus wrist data. *Physiological measurement*, *35*(11), 2183.
- Uswatte, G., Taub, E., Griffin, A., Rowe, J., Vogtle, L., & Barman, J. (2012). Pediatric Arm Function Test: reliability and validity for assessing more-affected arm motor capacity in children with cerebral palsy. *American journal of physical medicine & rehabilitation/Association of Academic Physiatrists*, *91*(12), 1060.
- Uswatte, G., Taub, E., Griffin, A., Vogtle, L., Rowe, J., & Barman, J. (2012). The pediatric motor activity log-revised: assessing real-world arm use in children with cerebral palsy. *Rehabilitation psychology*, *57*(2), 149.
- Wang, Q., Markopoulos, P., Yu, B., Chen, W., & Timmermans, A. (2017). Interactive wearable systems for upper body rehabilitation: a systematic review. *Journal of neuroengineering and rehabilitation*, *14*(1), 20.
- Whiting, P., Rutjes, A. W., Reitsma, J. B., Bossuyt, P. M., & Kleijnen, J. (2003). The development of QUADAS: a tool for the quality assessment of

studies of diagnostic accuracy included in systematic reviews. *BMC medical research methodology*, *3*(1), 25.

- Zoccolillo, L., Morelli, D., Cincotti, F., Muzzioli, L., Gobbetti, T., Paolucci, S., & Iosa, M. (2015). Video-game based therapy performed by children with cerebral palsy: a cross-over randomized controlled trial and a cross-sectional quantitative measure of physical activity. *Eur J Phys Rehabil Med*, *51*(6), 669-76.

4.3 Asymmetry of Upper Limb activity in children with unilateral cerebral palsy: validation of a Triaxial Accelerometer Approach

Bimanual upper limb activity is crucial for many activities of daily living, such as self-care, work, leisure, social communication. These activities encompass both gross and fine movements, and involve a tight form of cooperation between the upper limbs. The cooperation implies in the majority of cases by a distinction between right and left hand, since one conducts the action, whereas the contralateral helps in completing the motor task, playing a secondary role. For this reason, the two hands are commonly called "dominant" and "non dominant", respectively. In unilateral cerebral palsy (UCP) quantitatively measurement of the asymmetry in the use of upper limbs (ULs) could overcome the limitation of many outcome measures in which scores are dependent on the experience and training of the therapist. The primary purpose of this study was to determine the validity of accelerometry, by means of Actigraph GXT3+, worn on the both wrists, to measure asymmetry in use of the two upper limbs during the Assisting Hand Assessment in children and young aged 3-25 years old with Unilateral Cerebral Palsy (UCP) compared to age-matched typically developing (TD) subjects. The hypothesis to be tested is if triaxial accelerometers (Actigraph GXT3+), in a standardized setting, can discriminate the differences in the use of the two upper limbs in UCP subjects respect to TD and detect asymmetry in UCP group related to clinical outcome measures (Assisting Hand Assessment scores and MACS level).

4.3.1 Materials and methods

Study Participants & Setting

Data were collected at the Stella Maris Research Centre, Pisa - Italy and at the Queensland Cerebral Palsy and Rehabilitation Research Centre, Brisbane - Australia.

Potential participants were identified from hemiplegic children database of the Department of Developmental Neuroscience of the IRCCS Stella Maris (Pisa, Italy) for the UCP children and through a population-based research at the Queensland Cerebral Palsy and Rehabilitation Research Centre (Brisbane, Australia). A convenience sample of TD were identified as volunteers.

Eligible children and their parents were then invited to participate in the study trial and informed consent to participate was obtained from the child and/or by her/his parents prior to begin the assessment. Ethics approval was obtained from the various relevant ethics committees at participating hospitals and universities in Italy by Tuscany Paediatric Ethics Committee (78/2016) and in Queensland, Australia (NHMRC 465128). The study has been registered at the www.clinicaltrials.gov (NCT03054441).

Participants of the CP group were selected based on the following inclusion criteria:

- diagnosis of UCP
- age between 3 and 25 years old at study assessment
- located in Italy or Queensland (Australia)

Exclusion criteria were:

- medical complications that would interfere with study participation (e.g., uncontrolled seizures, epilepsy)
- dystonia or athetoid movements
- insufficient cognitive levels to comply with the instructions
- other progressive neurological disorders

- marked visual or hearing impairment.

Participants of the healthy group were selected based on the following inclusion criteria:

- age between 3 and 25 years old at study assessment
- no clinical documented disorders
- right-hand dominance
- located in Italy or Queensland (Australia)

Materials

- Actigraphs: Each participant wore an activity monitor (wGT3X-BT Monitor, ActiGraph, Florida, FL, model 7164; 4.6cm x 3.3cm x 1.5cm, 19g) on each wrist. The accelerometers were fastened to the wrist using custom-made Velcro wristbands. The Actigraph GT3X+® monitor was selected for this study as it has been identified as a reliable instrument for measuring movement intensity (Clanchy et al, 2011, Mitchell et al, 2015). The wrist was identified as potentially viable location for wearing the Actigraph GT3X+® monitor when assessing the use of upper limb in children.
- AHA: The Assisting Hand Assessment (β-version 5.0) which measures and describes the effectiveness with which a child with unilateral disability makes use of his impaired (assisting) hand during bimanual activities (Krumlinde-Sundholm et al, 2007). The AHA is scored from video recordings (lasting about 20 minutes) of the play activity, subsequently scored based on 20 predefined items using a four-point rating scale.
- There are different forms of the AHA, which are comparable with each other, allowing the comparisons amongst subjects of different ages and the detection of potential changes over time in the same subject (Holmefur et al 2007). Children aged 5-12 years old are tested with the School-Kids AHA, using a board game. There are two different themes

for the board games: “the Captive in the Fortress Game” and “the Alien Game”. In both versions, children have a mission to complete, which involves performing different tasks during the course of the game. On each board there is a path with several steps, linked to a specific card with the instruction about how to reach the next step. For children younger than 5 years old, or if greater than 5 years but are not interested in board game due to their cognitive level, toys from the AHA kit can be used as a free play session. For adolescent participants, the Ad-AHA board game “Go with the Ice Floe” is used. UCP and TD participants undertook different versions of the AHA assessment, depending on their age and on their cognitive ability.

- The Edinburgh Handedness Inventory is a simple and brief quantitative method of assessing of handedness, composed by 10 items (writing, drawing, throwing, scissors, toothbrush, knife without a fork, spoon, broom, striking a match, open the lid of a box) (Oldfield, 1971). Handedness is calculated based on the activities mainly done with right or left hand: it is the ratio between the difference of the two values divided by their sum, expressed in percentage. It can range from -100 (left-handed) to +100 (right-handed). Participants of TD group have been tested with Edinburgh Inventory.
- The Manual Ability Classification System (MACS) is a questionnaire which describes how children with CP use their hands to handle objects in daily activities. MACS describes five levels based on the children’s self-initiated ability to handle objects and their need for assistance or adaption to perform manual activities in everyday life. It is suitable for children between 4 and 18 years and the interpretation of the level is related to the age (Eliasson et al, 2006).

Methods

The trial was conducted in clinical environment, in a quiet room during a playful session.

Each child was asked to wear two wearable accelerometers (Actigraph G3XT+, preventively initialized and embedded on Velcro-strap bracelets) on both wrists and perform the Assisting Hand Assessment (AHA) test in the age-appropriate version: Kids-AHA for children between 18 months and 12 years (free play for children aged less than 5 years, alien game or fortress game for children aged between 5 and 12 years) or Ad-AHA board game "Go with the Ice Floe" for adolescents aged 12 years or more.

The AHA assessment was video recorded, as planned in the manual, and the start and end times of the test were registered.

In addition, MACS levels of UCP participants were rated by therapists together with participant's family. The Edinburgh Inventory Test (EHI) was performed as a structured interview to determine the handedness/laterality of each TD participant and it was scored with the online software http://zhanglab.wdfiles.com/local--files/survey/handedness.html .

AHA were scored by a certified AHA rater on the video recording and expressed in AHA units and Edinburgh Inventory

The Actigraph units sample at 80Hz which store summed values in random-access memory, which are subsequently downloaded to a personal computer. The data were processed with the Actigraph software. The number of accelerations was measured as average of activity counts and sum of activity counts, concerning Axis 1, Axis2, Axis 3 and Vector Magnitude. The was analysed with 1 second epochs, but the latter was also processed with 180 seconds epochs, in order to draw the related graphs, due the large amount of data.

The data were collected in 80Hz sample rate and in 1-sec epochs during AHA and analysed as a whole period. Raw data were summarized through Actilife Software 6 and expressed both as Average Counts, and Sum of Counts. Any

difference arose between the two methods, reason why in this paper only the average counts were reported. The Mean Activity was extracted separately for the dominant hand (DH) and non-dominant hand (NDH), regarding the values of Axis 1,2,3 and Vector Magnitude.

The asymmetry index was calculated processing the mean activity of each upper limb obtained by the accelerometry data of the AHA collection following the Edinburgh Inventory formula:

$$AI = \frac{MEAN\ ACTIVITY_{dominant\ hand} - MEAN\ ACTIVITY_{non\ dominant\ hand}}{MEAN\ ACTIVITY_{dominant\ hand} + MEAN\ ACTIVITY_{non\ dominant\ hand}} * 100$$

Clinical (AHA unit scores and MACS levels) and actigraphic (MADH, MANDH and AI) data were analyzed by means of the Statistical Package for Social Sciences (SPSS, version 20.0). Median and 95% confidence intervals (CI) were calculated and reported. Age and sex measures for both groups were calculated to check for differences between groups by means of Mann–Whitney U independent sample tests.

As a first step, intra group differences (for TD and for UCP) between the MADH and the MANDH were evaluated by means of Wilcoxon matched-pairs signed rank test. Between-group differences (UCP vs TD) for all actigraphic parameters (MADH, MANDH and AI) were evaluated, by means of Mann–Whitney U independent sample tests.

For the MACS level of UCP group, comparisons of three groups (MACS I, II and III) were made by Kruskall–Wallis. Two sided p-values <0.05 were regarded statistically significant and post hoc analyses were made by Mann Whitney U post hoc tests with p-values <0.017.

Finally, in the UCP group, Spearman rho correlation has been used to correlate the MADH, the MANDH and the AI with AHA scores, respectively.

4.3.2 Results

Participants

A total sample of 106 children (55 UCP and 51 TD) was evaluated and 100 of them met the inclusion criteria.
The reasons of exclusion were:

- Movement disorder associated with CP (n=2)
- Insufficient cognitive levels to comply with the instructions (n=3)
- Left-hand dominance for a TD subject (n=1)

The UCP group consisted of 50 subjects with CP (mean age 9,93±5,23 years, range 3-25) diagnosed as hemiplegia according to Hagberg's classification (Hagberg, Hagberg, & Olow, 1993). Gender and affected side were distributed as follow: male = 30, female = 20; and affected side was right = 33 and left = 17). 16 children had Manual Ability Classification System (MACS) level I, 23 had MACS II and 11 had MACS III (Eliasson et al., 2006).
The TD group consisted of 50 subjects (mean age 10,14±5,19 years, range 3-24,91, 30 Male, 20 Female). All TD presented complete right hand handedness, as confirmed by the Edinburgh Handedness Inventory scores (>0,8).
No statistically significant differences were found between the two groups for age and sex.

Intra and between groups comparisons of UCP and TD: Mean activity of dominant hand (MA_{DH}), Mean activity of non dominant hand (MA_{NDH}) and Asymmetry index (AI) (Table 1)

Within each group (UCP and TD) the MA_{DH} was significantly higher than the MA_{NDH} ($p<0.000001$ and $p<0.000001$, respectively).
For the between-groups comparisons, the MA_{DH} was significantly higher in the UCP group than the TD group ($p=0.001$) while the MA_{NDH} was significantly lower in UCP than in TD group ($p<0.000001$) (Fig.1).

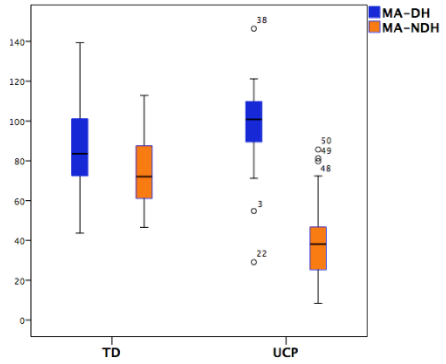


Fig. 1: Mean activity of DH and NDH in TD and UCP groups

Moreover, the values of AI of UCP were significantly higher than those of TD group ($p<0.000001$) (Fig.2).

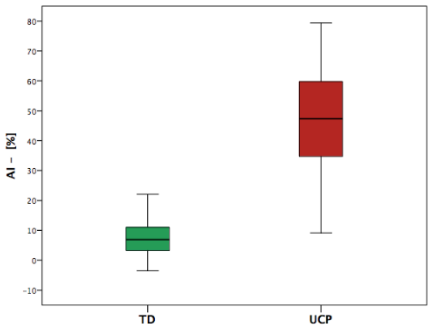


Fig. 2: Asymmetry index in TD and UCP groups

MEAN	UCP mean [CI 95%]	TD mean [CI 95%]	p*
MADH	97,71 [92,32-103,09]	86,60 [81,10-92,09]	0,001
MANDH	38,80 [33,65-43,95]	74,72 [70,10-79,33]	<0,000001
p**	<0,000001	<0,000001	
Asymmetry Index (AI)	45,88 [40,75-51,02]	7,27 [5,47-9,06]	<0,000001

** = Mann–Whitney U independent sample tests*

*** = Wilcoxon matched-pairs signed rank test*

Tab. 1: Mean activity of DH and NDH in TD and UCP groups

Relationship among MACS levels in UCP group

There was a significant difference among the MACS levels (I, II and III) and both clinical (AHA, $p<0.000001$) (Fig. 3) and actigraphic data (MA_{DH}, MA_{NDH} and AI; $p= 0.004$, 0.001 and <0.000001, respectively).

The AHA unit values were significantly higher for MACS I than MACS II ($p<0.000001$) and MACS III ($p<0.000001$) and also between MACS II and MACS III ($p<0.000001$).

The MA_{DH} values were significantly lower in the MACS I and in the MACS III than the MACS II ($p=0.004$ and $p=0.012$, respectively) and similar between MACS I and MACS III ($p=0.251$) (Fig. 4).

The MA_{NDH} values were significantly higher in the MACS I and in the MACS II than the MACS III ($p<0.000001$ and $p=0.001$, respectively) while the values were similar between MACS I and MACS II ($p=0.168$) (Fig. 5).

The AI values were significantly lower in the MACS I and the MACS II than the MACS III ($p<0.000001$ and $p=0.004$) and also in the MACS I respect to MACS II ($p=0.001$) (Fig 6).

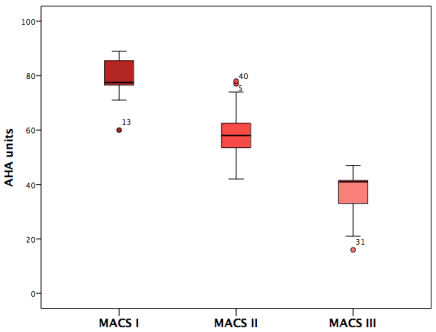


Fig.3: AHA scores of UCP children grouped by MACS levels levels

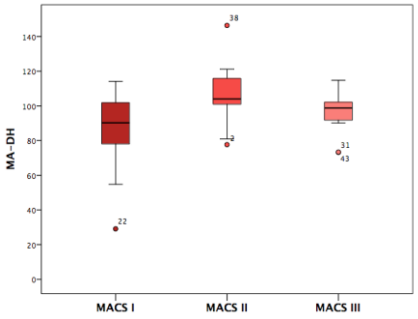


Fig. 4: Mean activity of DH of UCP children grouped by MACS

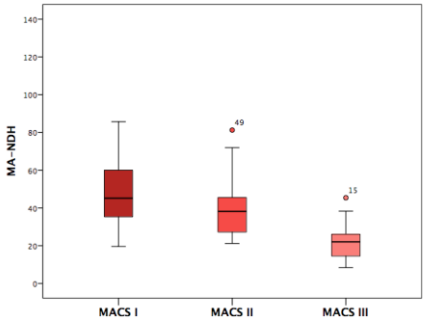


Fig. 5: Mean activity of NDH of UCP children grouped by MACS levels

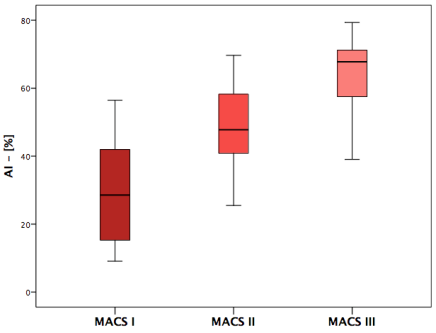


Fig. 6: Asymmetry index of UCP children grouped by MACS levels

UCP	MACS I (n. 16)	MACS II (n.23)	MACS III (n.11)	p*	p** I Vs II	p** I Vs III	p** II Vs III
MADH	86,90 ± 22,22	106,29 ± 15,01	95,47 ± 12,61	0,004	0,004	0,251	0,012
MANDH	48,33 ± 19,25	39,90 ± 14,91	22,62 ± 11,18	0,001	0,168	<0,000001	0,001
AI	29,89 ± 16,37	48, 70 ± 11,21	63,25 ± 12,60	<0,000001	0,001	<0,000001	0,004
AHA score	79,06 ± 7,59	58,70 ± 9,42	36,45 ± 10,38	<0,000001	<0,000001	<0,000001	<0,000001

** Kruskall–Wallis*

*** = post-hoc analysis by Mann–Whitney U independent sample tests*

Tab. 1: clinical and actigraphic parameters in MACS levels of UCP group

AHA Correlations in UCP group

Scores of UCP group ranged from a minimum of 16 to a maximum of 89 AHA units.

There was a significant negative correlation between AHA unit scores and the MA_{DH} values (Spearman' Rho= -0.285, p=0.045) (Fig. 7) and the AI (Spearman' Rho= -0.819, p<0.000001) (Fig. 8) while it was significantly positive between AHA unit scores and the MA_{NDH} (Spearman' Rho= 0.668, p<0.000001) (Fig. 7).

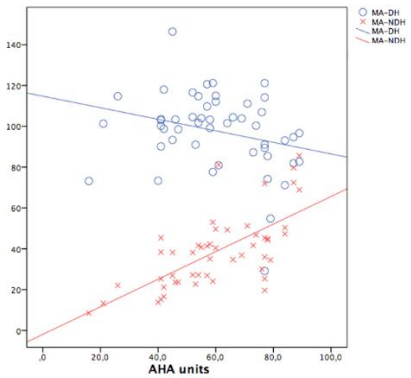


Fig. 7: Correlation between mean activity of DH and NDH and AHA scores in UCP group

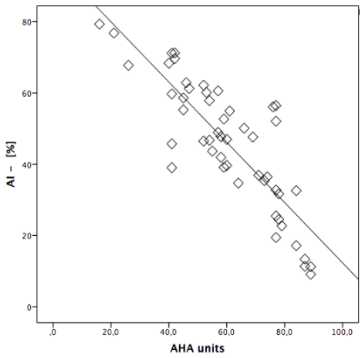


Fig. 8: Correlation between asymmetry index and AHA scores in UCP group

4.3.3 Discussion

This is the first study aimed to evaluate the validity of accelerometry (Actigraph GXT3+) for detecting the asymmetry in the use of the two upper limbs in children with UCP. The semi-structured setting of the Assisting Hand Assessment allowed us to have a reliable environment, age-related, in which the spontaneous use of the two upper limbs can be detected both in TD and UCP subjects. The use of the actigraphs for each upper limb was the only additional request respect to the traditional assessment with the AHA. Performing the AHA wearing the two actigraphs did not interfere with subjects' playing and they were well accepted due to their resemblance to any digital watch.

The first important result is that the data coming from the actigraphs were able to quantify the differences in the mean activity between the two hands (DH and NDH) in both groups (TD and UCP) showing that the DH was used more than NDH. The DH, in fact, also in TD, has the more active role during bimanual activities than the NDH that makes functions as support and stabilization, so it is expected a lower quantity of movement. However, in TD the asymmetry index was very low demonstrating a high cooperation between the two hands in contrast with significantly higher values of asymmetry in UCP. It has to be considered, in fact, that manual activities typically require the co-operation of both hands, which tend to be specialized for different functions (Guiard, 1987; MacNeilage, 1990). For instance, when we remove the lid from a jar or we button up a shirt, the NDH holds the object in a stable position while the DH acts upon it. Hence, the NDH plays a postural role in stabilizing the grasped object and at the same time provides a spatial reference frame into which the DH manipulates the object. However, saying that the NDH offers stability does not mean that the hand is immobile. On the contrary, the NDH ensures a static or dynamic stabilization. For instance, in handwriting, the pen cannot be dexterously manipulated by the DH if the page is not stabilized and

periodically re-positioned by the NDH so that the position and orientation of the page always remain appropriate to the DH action. Based on these previous findings, our results confirm that in TD, even if the NDH have lower values than the DH, merging together there is not an important asymmetry.

By the comparison of mean activity of DH and NDH in TD vs. UCP children, we found out that the mean activity of DH of UCP was higher than in TD. This finding could be in contrast with other studies (Cooper et al, 1995; Arnould et al, 2006; Duquè et al, 2003, Rich et al, 2017), in which there is reported that the nonparetic hand of children with hemiplegia was significantly impaired, although to a lesser extent than the paretic hand, respect to the dominant hand of TD groups. However, the actigraphs data are referred to the quantity of movement and the higher activity could be related to the higher movements that the UCP subject need to perform with his DH in order to supply the activity of affected hand.

For the UCP group, the data confirm that children with UCP have developed their handedness on the less affected side (Arnould et al, 2007). The values of MA of DH and of NDH showed some outliers in UCP group. In particular, for the MA of DH there was the subject #38 who had a very high value and it could be related to the particular active behaviour of the child during the assessment, while subjects #3 and #22 are were young adults who were very quiet and they moved their non affected arms few times. The three outliers (#48, #49 and #50) with higher values of MANDH were the subjects with higher values of AHA.

Another important finding was that there were differences in actigraphic data among the MACS levels. MACS was designed to describe how children with CP use their hands to handle objects in daily activities and it is expected that it emerges also during AHA, which evaluates the spontaneous use of the assisting hand (affected hand) during a semi-structured play session requiring bimanual handling. All the three classes of MACS (I, II and III) had significantly different AHA scores according to literature data (Nordstrand et al, 2016). However, there were few outliers, the case #13 and the case #31 for lower values than expected for their MACS level I and III, respectively. On the

contrary cases #5 and #40 had higher values than expected for their MACS level II. These few discrepancies could be related to a lower or higher performance during the assessment as AHA respect to the use of their hands during the daily activities.

By the deep analysis of the mean activity of each hand separately in relation to MACS levels we found a significant difference among them, but while for the MADH it was mainly related to the values of MACS level II that were higher than level I and III, for the MANDH the MACS level III had lower values than the two other groups (MACS I and MACS II). This finding could be related to the crucial difference between MACS II and MACS III. MACS II can achieve required tasks using alternative ways of performance with both hands and also extra actions of DH for have success in bimanual tasks that can determine higher values of MADH. It usually does not happen in MACS III who required environmental supportiveness for achieving goals. When comparing MACS levels with AI, the data confirmed that there were significant differences among levels with a significant raising of the AI so that the UCP children at MACS level III have the highest level of asymmetry. All the findings in UCP children were confirmed by the correlation between AHA unit scores and all the quantitative actigraphyc parameters (MADH and MANDH and AI). As expected, higher AHA score correspond to a good symmetry of upper limbs use (and therefore low AI values) and the lower activity of DH with higher values of NDH.

Limitations

Three aspects of the analysis need a further analysis. Firstly, the role of mirror movements has to be deepened because a systematic clinical assessment of quantity and quality of mirror movements is useful for a more complete evaluation of the characteristics of upper limb movements and their correlation to the MA of each hand.

4.4 Final remarks

Understanding the development of bimanual upper activity both in children with typical development and with neurodevelopmental disorders is important, however such knowledge requires further development of quantitative tools such as wearable sensors. The systematic review summarizes the growing body of literature concerning the available clinical application of inertial sensors worn on both upper limbs in typically developed children and peers with neurodevelopmental disorders. Existent studies support the use of accelerometers as a potential future reliable, valid and objective outcome measure, appropriate to measure upper limb activity in young children with neurodevelopmental disorders, to determine intervention effectiveness, due to its high interrater reliability and strong concurrent validity with validated measures of capacity; however, there are many limitations emerged, that are small samples sizes, lack of control groups, variety of Actigraphs types and parameters used.

The work of validation of the Triaxial Accelerometer Approach with Actigraphs, could be an important milestone from which start for a better and deeper knowledge of upper limb movements by an easy assessment which could be added to the traditional evaluation without require time or effort to the child. Furthermore, the possibility to have access to quantitative data suggests the use of wearable sensors also for planning upper limb activity strategies in interventions, making them more customized in relation to subjects' rehabilitative needs and, finally, it provides the opportunity to have a high sensitive tool for detect spontaneous and intervention-induced changes.

Data reported in this Chapter are based on the following paper:

- Braito I, Maselli M, Sgandurra G, Inguaggiato E, **Beani E**, Cecchi F, Cioni G, Boyd R, Assessment of upper limb use in children with typical development and with neurodevelopmental disorders by inertial sensors: a systematic review, *submitted*
- **Beani E**, Maselli M, Sgandurra G, Sicola E, Perazza S, Cecchi F, Dario P, Boyd R, Cioni G Asymmetry of Upper Limb activity in children with unilateral cerebral palsy: validation of a Triaxial Accelerometer Approach, *submitted*

CHAPTER 5

ACTION-OBSERVATION THERAPY (AOT) AND INFORMATION AND COMMUNICATIONS TECHNOLOGIES (ICT) FOR HOME REHABILITATION OF CHILDREN WITH HEMIPLEGIA

5.1 Action Observation Training for Rehabilitation in Stroke: a Systematic Review and Meta-analysis

Action observation therapy (AOT) is an innovative method of rehabilitation for stroke patients which, thanks to the mirror neuron system, improves the patient's limb function. Many studies on the upper and lower limb in adults, children and adolescents were conducted using this rehabilitation tool. This systematic review and meta-analysis was carried out using six different databases, where 23 articles that matched the inclusion criteria were identified. The paper discusses the use of AOT for rehabilitation in stroke patients and compares its use on the upper and lower limb function in adults and children. The efficacy of AOT according to the ICF domains, the length of the training, the type of training performed by the patients and the setting are analysed. This systematic review suggests that AOT has a positive effect on the patient's limb function. The authors make recommendations for future studies which are needed due to the lack of uniformity of the present literature.

5.1.1 Methods

Search strategy

A literature search was conducted using six electronic databases: PubMed, EBSCO, Cochrane, Scopus, Web of Science and Eric. These databases were selected as they represent a broad spectrum of disciplines that perform clinical research.

The research was carried out from 1 February 2016 to 14 September 2017 using the terms: ((\"action observation\") AND (stroke OR hemipleg* OR (\"cerebral palsy\")) AND (intervention OR treatment OR trial OR training).

The search was limited to articles or reviews written in English and published after 2000. It was not possible to use these filters in the Cochrane library and in the Eric electronic database.

The year limit was chosen because the MNS was only discovered in the 1990s by Rizzolatti et al. (Rizzolatti et al., 1996a) and in 1995 the first article on the MNS in human was published by Fadiga et al (Fadiga et al., 1995).

Inclusion and exclusion criteria

The inclusion criteria for the systematic review were articles that:

1. had been published after 2000
2. had an AOT training on upper and/or lower limb
3. were carried out on patients with stroke or CP
4. were carried out in children or adults
5. were fully written in English.

The exclusion criteria were:

1. studies that included AOT training with total length shorter than one week
2. studies that included only observation sessions not followed by the action
3. studies on patients with other diseases such as Parkinson's disease, orthopaedic patients or aphasia.
4. abstracts, reviews, theses or conference papers.

Methodological quality

Reviewers screened both the title and abstract of all the potentially relevant publications and the duplicates were removed. Each reviewer then examined the abstracts selected and chose the articles based on the inclusion and exclusion criteria. Following completion of the searches, all publications were fully read by authors. Disagreements on whether a study met the inclusion criteria, level of evidence, or quality ratings were resolved by consensus.

The reviewers independently assessed the trials and independently documented the methodological quality of the trials using the last version (March 2009) of Oxford Centre for Evidence-based Medicine – Levels of Evidence (CEBM),

and for RCT studies also the Physiotherapy Evidence Database (PEDro) scale (PEDro).

In the CEBM five possible levels are given. The levels used where those regarding therapy and are the following:

Level 1a: systematic review (with homogeneity) of RCTs

Level 1b: Individual RCT (with narrow Confidence Interval)

Level 1c All or none

Level 2a SR (with homogeneity) of cohort studies

Level 2b Individual cohort study (including low quality RCT; e.g., <80% follow-up)

Level 2c "Outcomes" Research; Ecological studies

Level 3a SR (with homogeneity) of case-control studies

Level 3b Individual Case-Control Study

Level 4 Case-series (and poor quality cohort and case-control studies)

Level 5 Expert opinion without explicit critical appraisal, or based on physiology, bench research or "first principles")

Criteria of homogeneity and of poor quality studies are further specified in the CEBM.

The PEDro scale was calculated on a score of 10 points: the first criterion was excluded as the guidelines of the PEDro scale on the web recommend. If the score assigned by the reviewers was different, an agreement was found.

Data extraction and Outcome analysis

Before starting the data extraction we firstly agreed in pointing out specific research questions to answer that were:

1- Did the studies focus on patients with upper limb or lower limb dysfunction?
2- Are the studies on AOT focused on children and/or adults?
3- What type of patients were selected for the studies?
4- What kind of AOT training was conducted?
5- How long did the treatment last?

6- Where did the training sessions take place? Was the treatment home-based?
7- What were the main outcome measures used?

Based on these questions, the selected papers were categorized into two groups according to the body section the AOT treatment focused on: upper limb and/or the lower limb.

Subsequently, the two groups were further divided into two subgroups: those dealing with adults and those dealing with children and adolescents (less than 18 years old).

For each study we described the study design, the CEBM LEVEL, the PEDro score, the study aim, the diagnosis, the sample size (specifying the number of participants and their mean age), the setting, the duration and the intensity of the training. Moreover, the type of AOT performed, the perspectives of the videos and other treatments carried out in the experimental group were analysed. The type of intervention and other current treatments were specified for the control group. Finally, we analysed the study outcome measures, the results and the findings of all the study examined. These items can all be find in tables (n…).

Furthermore, the outcome measures of the randomized studies regarding the upper limb and the lower limb were divided into ICF domains. In those cases where an outcome measure covered more than one domain (such as the MUUL with items covering the body function and other the activity) we assigned the outcome measure to the most representative domain (Hoare et al., 2011). A special program ProMeta3, was chosen to verify the efficacy of AOT in the three ICF domains (body function, activity and participation). This program allowed us to calculate the effect sizes from different types of data in order to combine them even when the studies had different statistic measures and to calculate the overall effect size.

According to Cohen (Cohen, 1977), values of effect sizes between 0.2 and 0.5 were considered “small”, between 0.5 and 0.8 “medium”, and > 0.8 “large”.

5.1.2 Results

The database search identified a total of 326 articles. After removing duplicates a total of 144 records were selected. Based on title and abstract, 33 matched the inclusion criteria. After reading the full text, 23 were included in this review. Details can be seen in the flow chart in Fig. 1. Ten papers were excluded for various reasons: i) four papers were based on AOT training of less than one week; ii) one was a thesis; iii) three were not written in English,; iv) one was merely an abstract and v) the last one was based only on observation with no execution phase.

General findings

The first article regarding AOT that matched the inclusion criteria for this review was published in 2007 (Ertelt et al., 2007), and one in 2010 (Franceschini et al., 2010). From 2012 at least two articles per year were published, but most of the articles were published in 2013 (Cowles et al., 2013; Kim et al., 2013; Kim and Lee, 2013; Lee et al., 2013; Sgandurra et al., 2013) and in 2015 (Kim and Kim, 2015a, b; Park and Hwangbo, 2015; Sugg et al., 2015; Zhu et al., 2015). Of the 23 studies selected 13 were carried out in Korea (Bae and Kim, 2017; Kim and Bang, 2016; Kim and Kim, 2015a, b; Kim and Lee, 2017; Kim et al., 2013; Kim and Lee, 2013; Kim et al., 2014; Lee et al., 2013; Lee et al., 2017; Park and Hwangbo, 2015; Park et al., 2017; Park et al., 2014), 5 were performed in Italy, (Buccino et al., 2012; Franceschini et al., 2010; Franceschini et al., 2012; Sale et al., 2014; Sgandurra et al., 2013) 2 in the UK (Cowles et al., 2013; Kirkpatrick et al., 2016) and the remaining 3 were studies performed in Australia(Sugg et al., 2015), Germany (Ertelt et al., 2007), China (Zhu et al., 2015).

15 of the above mentioned studies focused on upper limb function, (Kirkpatrick et al., 2016) (Kim and Bang, 2016) (Kim and Kim, 2015a; Zhu et al., 2015) (Kim and Kim, 2015b) (Sugg et al., 2015) (Kim et al., 2014) (Sale et

al., 2014) (Lee et al., 2013) (Sgandurra et al., 2013) (Cowles et al., 2013) (Buccino et al., 2012) (Franceschini et al., 2012) (Franceschini et al., 2010) (Ertelt et al., 2007) and 8 on the lower limb (Kim and Lee, 2017) (Bae and Kim, 2017) (Park et al., 2017) (Lee et al., 2017) (Park and Hwangbo, 2015) (Park et al., 2014) (Kim et al., 2013) (Kim and Lee, 2013) .

Four of the 23 studies were carried out on children with CP (Kirkpatrick et al., 2016) (Kim et al., 2014) (Sgandurra et al., 2013) (Buccino et al., 2012) while the remaining studies were conducted on adults. In two of the four studies carried out on children with CP all the children had Unilateral CP (UCP) (Sgandurra et al., 2013) (Kirkpatrick et al., 2016); in one study (Kim et al., 2014) some children had UCP, while other children had bilateral CP while, in a further study (Buccino et al., 2012) most of the patients had UCP and few patients had bilateral CP.

Of the 19 studies on adults 5 studies included patients with sub-acute stroke whereas the remaining were on patients with chronic stroke (Cowles et al., 2013; Franceschini et al., 2012; Kim and Bang, 2016; Sale et al., 2014; Zhu et al., 2015).

The sample size for all the studies varied from a minimum of 12 people (Kim and Kim, 2015a, b) to a maximum of 102 (Franceschini et al., 2012) . The AOT performed consisted in a variety of approaches depending on the study reviewed. In some cases video performing the actions were shown to the experimental group (Bae and Kim, 2017; Buccino et al., 2012; Ertelt et al., 2007; Franceschini et al., 2010; Franceschini et al., 2012; Kim and Bang, 2016; Kim and Lee, 2017; Kim et al., 2013; Kim and Lee, 2013; Kim et al., 2014; Lee et al., 2013; Lee et al., 2017; Park et al., 2017; Park et al., 2014; Sale et al., 2014; Sgandurra et al., 2013; Sugg et al., 2015; Zhu et al., 2015); in others the therapist or mother (in a home-based study) performed the action (Cowles et al., 2013; Kirkpatrick et al., 2016).

In most of the studies the control group watched videos where no action was shown (Buccino et al., 2012; Ertelt et al., 2007; Franceschini et al., 2010; Franceschini et al., 2012; Kim and Lee, 2017; Kim et al., 2014; Park and

Hwangbo, 2015; Park et al., 2017; Park et al., 2014; Sale et al., 2014; Sgandurra et al., 2013; Sugg et al., 2015); while in other studies an action was performed without an observation phase (Cowles et al., 2013; Kim and Bang, 2016; Kim and Kim, 2015a, b; Kirkpatrick et al., 2016; Lee et al., 2013). In one study the patients in the control group underwent routine rehabilitation treatment. (Zhu et al., 2015).

The length of the AOT varied from two weeks (Sugg et al., 2015) to 12 weeks (Kirkpatrick et al., 2016), and from a minimum of 10 minutes per session a day (Lee et al., 2013) to a maximum of 90 minutes per session a day (Ertelt et al., 2007).

The average number of weeks was 4,78 ± 2,19 whereas the average length of the total time of the treatment session was 36,09 ± 20,05 minutes.

2 of the 23 articles selected were home-based (Kirkpatrick et al., 2016) (Sugg et al., 2015); 20 were held in a laboratory or in an in-patient hospital, whereas the setting was not specified in one article (Bae and Kim, 2017).

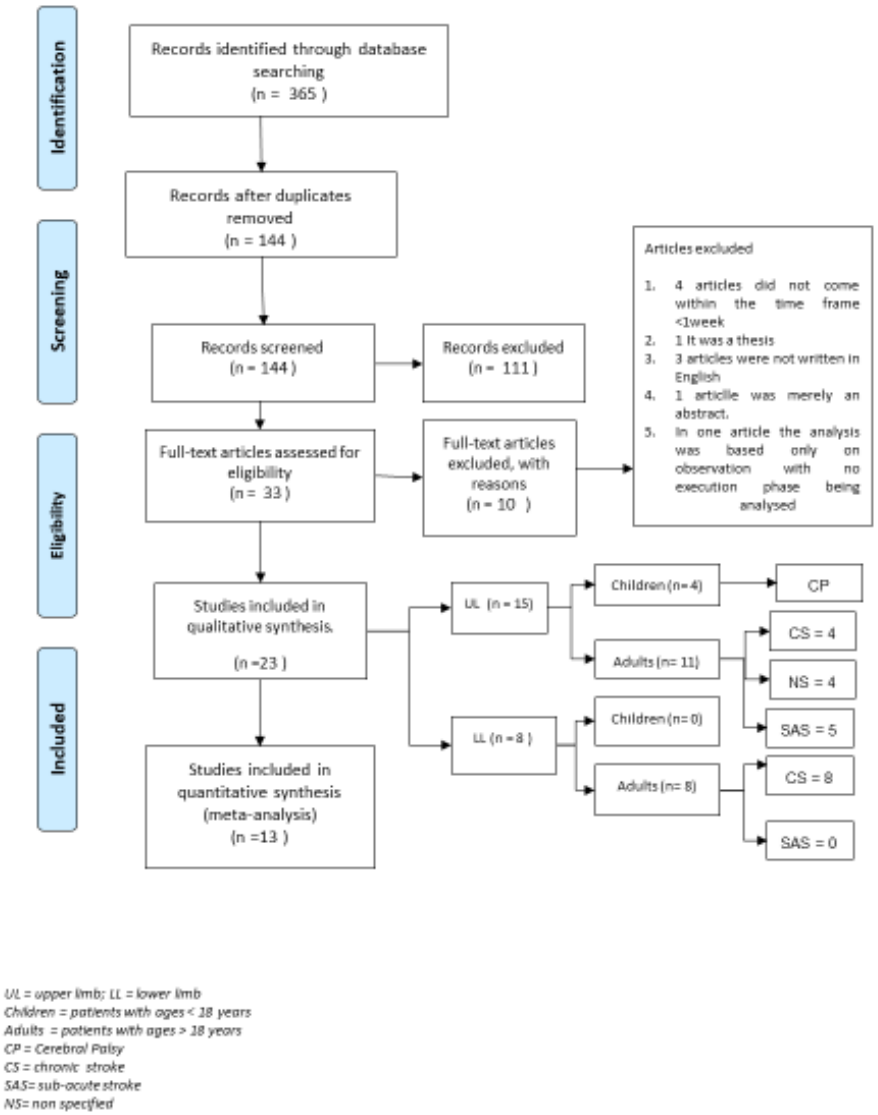
All the ICF domains were assessed across the different studies but each domain was explored using different outcome measures. Only in few studies the body function and activity domains were assessed using the same outcome measures (e.g. FMA, Melbourne, BBT, AHA)

The CEBM level was applied in all studies and most of them were classified as 2(Buccino et al., 2012; Cowles et al., 2013; Kim and Bang, 2016; Kim and Kim, 2015a, b; Kim et al., 2013; Kim and Lee, 2013; Kim et al., 2014; Lee et al., 2017; Park and Hwangbo, 2015; Park et al., 2014; Zhu et al., 2015), 9 studies resulted 1b (Bae and Kim, 2017; Ertelt et al., 2007; Franceschini et al., 2012; Kim and Lee, 2017; Kirkpatrick et al., 2016; Lee et al., 2013; Park et al., 2017; Sale et al., 2014; Sgandurra et al., 2013) and only two (Franceschini et al., 2010; Sugg et al., 2015)were CEBM LEVEL 4.

The PEDro scale was applicable in 21 studies and the results are shown in Table 5, 6, 7. Most studies(Bae and Kim, 2017; Buccino et al., 2012; Cowles et al., 2013; Franceschini et al., 2012; Park et al., 2017; Park et al., 2014; Zhu et al., 2015) obtained 7/10, only one study scored 9/10 (Kim and Bang, 2016).

Many articles offer preliminary evidence in support of the effectiveness of AOT with physical practice but many did not find significant results when AOT and placebo treatments were compared. Of all the studies examined, 17 out of 23 (74%) showed significant results in at least one outcome measure (Bae and Kim, 2017; Buccino et al., 2012; Ertelt et al., 2007; Franceschini et al., 2010; Franceschini et al., 2012; Kim and Bang, 2016; Kim and Kim, 2015a; Kim and Lee, 2017; Kim et al., 2014; Lee et al., 2013; Park and Hwangbo, 2015; Park et al., 2017; Park et al., 2014; Sale et al., 2014; Sgandurra et al., 2013; Sugg et al., 2015; Zhu et al., 2015).

With regard to the body level, all four studies carried out in children were focused on upper limb while out of the 19 studies in adults 11 studies focused on upper limb. In the following paragraphs we have separately analysed the studies on upper and the lower limb function in children and adolescents and in adults.



UL = upper limb; LL = lower limb
Children = patients with ages < 18 years
Adults = patients with ages > 18 years
CP = Cerebral Palsy
CS = chronic stroke
SAS= sub-acute stroke
NS= non specified

Fig. 1 Flow chart

Studies focused on upper limb in children and adolescents

Sample participants

All the studies included children with cerebral palsy (CP), with a cognitive level within normal limits and who did not have sensory impairments (Buccino

et al., 2012; Kim et al., 2014; Kirkpatrick et al., 2016; Sgandurra et al., 2013). As written before, not all the studies included children with same type of CP (i.e. hemiplegia) and in addition the aetiology, such as stroke, was not specified in any of the studies.

Two of the four studies excluded children who had had previous interventions: in one study patients were not included if there had been an intervention in the previous 3 months; such intervention included upper limb botulinum toxin injections, surgical intervention, or introduction of Lycra sleeves/suits or similar (Kirkpatrick et al., 2016). In the other study (Sgandurra et al., 2013) previous orthopaedic surgery or a botulinum toxin injection in the upper limb within 6 months prior to the study entry was an exclusion criteria.

In the study by Sgandurra et al. (Sgandurra et al., 2013) specific motor conditions were required for inclusion: a grade ≤2 on the Modified Ashworth Scale (MAS) and a grade between 4 and 8 on the House Functional Classification System (HFCS). In the study by Kim et al (Kim et al., 2014) children were included if the grade of MAS did not exceed grade 3 and with upper limb muscle strength not lower than A+ in manual muscle tests.

Epilepsy was an exclusion criterion: the participant did not take antiepileptic treatment (Buccino et al., 2012) or had no history of seizures or seizures well controlled by therapy (Sgandurra et al., 2013).

AOT training and control condition

In three articles (Buccino et al., 2012; Kim et al., 2014; Sgandurra et al., 2013), the AOT was performed by watching a video, while in the study by Kirkpatrick AOT consisted in watching a parent performing the action (Kirkpatrick et al., 2016). The content of the videos were actions related to the children's daily lifes (Buccino et al., 2012; Kim et al., 2014). In the study by Sgandurra (Sgandurra et al., 2013) the video showed unimanual or bimanual goal-directed actions. The specific actions performed are well described in the paper. There were different action performed and these were divided into levels of difficulty in two studies (Kim et al., 2014; Sgandurra et al., 2013); in Sgandurra et al.

2013 (Sgandurra et al., 2013) there were two different sets conceived in relation to the HFCS level and in each set the complexity increased during the sessions.

The perspective the video was not specified in one study (Buccino et al., 2012) even though it was stated that the videos were shown from different perspectives.

In the home-based study (Kirkpatrick et al., 2016) the parent performed the action while sitting next to the child on the side of the less-affected hand and facing in the same direction, so that the child observed the hand movements from an egocentric viewpoint.In one study (Kim et al., 2014) the patient observed the action on videos from the front, sides and back so that the subjects could observe the action in three dimensions, while in another study the action was presented only from a first-person perspective (Sgandurra et al., 2013). It is therefore difficult to understand how much the perspective influences the outcome and if an ideal perspective exists. The children in the control group in the studies that had used a video were asked to observe some photos (Kim et al., 2014), play computer games (Sgandurra et al., 2013) or watch video clips with no motor specific content (Buccino et al., 2012) and were then asked to perform the action. The children in the control group in the study that had not used videos, were asked to perform the action without observing the parent's hand movement (Kirkpatrick et al., 2016). The type of treatment suggested to the AOT control group was reasonably similar because three of the four studies (Buccino et al., 2012; Kim et al., 2014; Sgandurra et al., 2013) showed pictures or videos? with no motor content followed by the action required to the experimental group.

Training length

The length of treatment time varied considerably.

In one study (Kirkpatrick et al., 2016) the training was performed five days a week for 12 weeks for 15 min each session. In two studies, training was done five days a week for 3 weeks, one for 60 min each session (Sgandurra et al.,

2013), while the length of each session was not specified in the other study (Buccino et al., 2012). In the last study (Kim et al., 2014) the training was done three days a week for 4 weeks for 30 min each session.
Even though the studies had different lengths the total time of training of two of the four studies was 15 hours (Kirkpatrick et al., 2016; Sgandurra et al., 2013).

Setting

Three of the four studies on children were performed mainly in a laboratory or in an in-patient hospital, only one was home-based (Kirkpatrick et al., 2016). The latter study did not demonstrate a significant improvement in hand function between the control and the experimental group, while in the Sugg et al. study, carried out on adults, did (Sugg et al., 2015). Further studies are necessary to understand if the home setting can successful be used in AOT rehabilitation

Outcome measures.

Many different outcome measures were used. In the Fig. 2, the various outcome measures are shown according to their ICF domains.
All four studies analysed the body function ICF domain using Melbourne Assessment (Buccino et al., 2012; Kirkpatrick et al., 2016; Sgandurra et al., 2013) or Grasp Power and Modified Ashworth Scale (MAS) (Kim et al., 2014)
Three of the four studies had at least one outcome measure of the ICF activity domain (AHA (Kirkpatrick et al., 2016; Sgandurra et al., 2013), ABILHAND-kids (Kim et al., 2014; Kirkpatrick et al., 2016; Sgandurra et al., 2013) or BBT(Kim et al., 2014)). In the ICF activity domain, ABILHAND-kids did not report any statistically significant difference in all the studies.
Only one study investigated the ICF Participation domain with the WEE-FIM test (Kim et al., 2014). The latter did not record any statistically significant difference between groups.

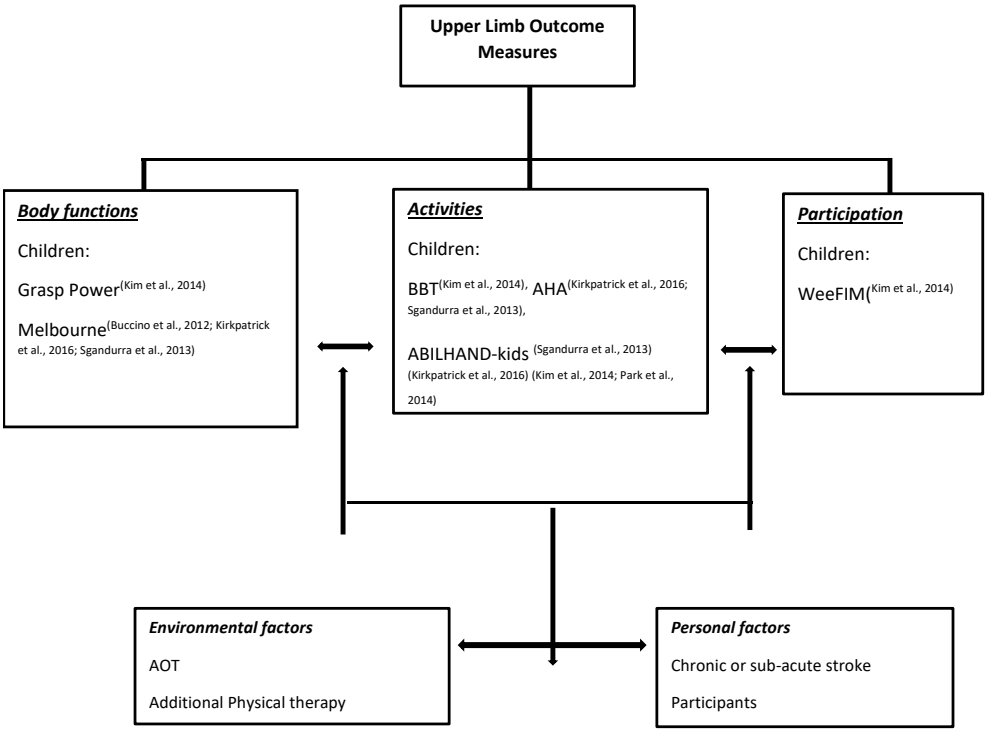


Fig.2

Effectiveness of AOT

Three out of the four studies showed a significant difference between groups (Buccino et al., 2012; Kim et al., 2014; Sgandurra et al., 2013) while one (Kirkpatrick et al., 2016) did not. Even though the same outcome measures of Sgandurra et al.(Sgandurra et al., 2013) were used, as reported in the paper, key differences between trials included setting, dose, duration of the therapy and sample size. The effectiveness was both reported in in the body function using MUUL (Buccino et al., 2012) and in the activity domain using BBT (Kim et al., 2014) and AHA (Sgandurra et al., 2013). However, in the study Sgandurra et al. (Sgandurra et al., 2013) MUUL was used as a secondary outcome measure but differences between groups were not significant.

Studies focused on upper limb in adults

Sample participants

Studies on adult upper limb included a very heterogeneous sample of participants. Four studies included chronic stroke (>6 months duration) (Ertelt et al., 2007; Franceschini et al., 2010; Lee et al., 2013; Sugg et al., 2015); two enrolled only patients with first-ever stroke 30 days (±7) after the onset of the event with ischemia or primary haemorrhage (Franceschini et al., 2012; Sale et al., 2014); one (Cowles et al., 2013) included adults who had suffered a stroke between 3 and 31 days before recruitment with an intact premotor area; and the remaining two studies (Kim and Bang, 2016; Zhu et al., 2015) involved subjects within 6 months of stroke.

Two studies did not specify if the patients were affected by subacute or chronic stroke (Kim and Kim, 2015a, b), their age, or their clinical conditions.

The subjects enrolled had to be able to understand and perform given instructions, and cognitive assessment was evaluated by Mini-Mental State Examination score (MMSE) (Franceschini et al., 2012; Kim and Bang, 2016; Lee et al., 2013; Zhu et al., 2015), by Montreal Cognitive Assessment (MoCA) (Sugg et al., 2015), or by Token test (Franceschini et al., 2010; Franceschini et al., 2012). Impaired comprehension or dementia were considered exclusion criteria (Ertelt et al., 2007; Sale et al., 2014).

Some studies specified various motor inclusion criteria (Cowles et al., 2013; Sugg et al., 2015; Zhu et al., 2015) such as being able to produce movement in the paretic upper limb.

AOT training and Control condition

Videos showing different actions (Kim and Bang, 2016; Lee et al., 2013; Sugg et al., 2015; Zhu et al., 2015) or videos of daily routines (Ertelt et al., 2007; Franceschini et al., 2010; Franceschini et al., 2012; Sale et al., 2014) were used in the the experimental group.

In three studies (Franceschini et al., 2010; Franceschini et al., 2012; Sale et al., 2014) the actions were daily routine unimanual and bimanual actions e.g. drinking from a glass, combing hair and were divided into three acts.

In a study (Sugg et al., 2015) the actions performed were not specified. In another study (Lee et al., 2013) the study task was the action of stretching out the right hand to pick up a cup, bring the cup to the mouth to touch the mouth and then return the cup to its initial position. In the article by Zhu et al; (Zhu et al., 2015) the video showed both simple movements e.g. bending and extension, abduction and adduction, pronation and supination and complex movements e.g. the catching and the releasing of small balls, handling a pen. In the study by Ertelt et al. (Ertelt et al., 2007) the type of action performed was not specified but it was said that both hands had to be used to complete the action. In a further study (Cowles et al., 2013) the action observed was performed by a therapist who sat alongside the participant. The therapist performed unimanual actions, which were functional activities such as. bring a telephone to the ear or pour water. The activities were changed if the participant improved or got bored. In two studies (Kim and Kim, 2015a, b), it was not clear what type of AOT treatment was used. However, both studies were based on Feys et al study,1998 and the actions performed included various activities i.e. feeding, grasping a can, combing hair. In four other articles a video of daily routine actions was shown to the patients. As the aim of the rehabilitation programme is to acquire the ability to perform some routine actions this type of treatment should help to achieve that aim.

In four studies (Ertelt et al., 2007; Franceschini et al., 2010; Sale et al., 2014; Zhu et al., 2015) the difficulty of the actions proposed in the videos increased during the treatment. In the article by Zhu et al. (Zhu et al., 2015) the complexity of the action increased during each session also.

The model performing the action (Kim and Bang, 2016) was a healthy woman, while in three studies (Franceschini et al., 2010; Franceschini et al., 2012; Sale et al., 2014) the models were non-disabled people, either men or women,

different from video to video and in another study (Lee et al., 2013) the model was an adult male.

In some studies, the perspective was specified. In one study (Kim and Bang, 2016) three perspectives were provided simultaneously: front, side and top, in another (Zhu et al., 2015) the action was seen from straight on, right above and right inside whereas in a further study (Lee et al., 2013) the video was shot from the front. In two studies (Franceschini et al., 2012; Sale et al., 2014) the actions were observed from a first-person perspective, two studies (Ertelt et al., 2007; Franceschini et al., 2010) reported that the action had been recorded from different perspectives (one (Ertelt et al., 2007) specified that three perspectives had been used) but failed to mention them, and in three others no mention was made at all of the type or number of perspectives.(Kim and Kim, 2015a, b; Sugg et al., 2015)

In one of the studies (Cowles et al., 2013) which showed the therapist sitting at a table alongside the participant on their paretic side, the therapist used the upper limb that matched that of the participant's paretic side to demonstrate the action: hence both were on the same plane. Similarly to the studies on children, the use of different perspectives does not allow us to understand whether some perspectives are better than others or to assess whether the type of perspective used is relevant or not. Likewise in the study on children where the parent performed the action (Kirkpatrick et al., 2016), no significant difference was found. This finding could be related to the lack of standardization of the perspective used or to other characteristics of the study.

The control group performed the action without observing it (Kim and Bang, 2016; Kim and Kim, 2015a, b), or they observed videos, images or sequences of geometric symbols (Ertelt et al., 2007) which showed a neutral environment (Franceschini et al., 2012; Sale et al., 2014; Sugg et al., 2015) and performed the same actions as the experimental group.

In one particular study (Franceschini et al., 2012) the patient after watching static images was asked to perform limb movements, as well as possible, for 2

minutes following a standard sequence, and simulating those performed by the experimental group which involved shoulder and elbow joint mobilization.

In two studies the patient underwent the same therapy as the experimental group, apart from AOT (Cowles et al., 2013; Zhu et al., 2015). There were two control groups in one other study: one which only performed the action while the other control group neither watched the video nor practised the actions (Lee et al., 2013). In another study the control group was composed of the same patients at the baseline (Franceschini et al., 2010).

Training length

The studies on adults were carried out for a different length of time: one study (Sugg et al., 2015) lasted only two weeks, two of them (Cowles et al., 2013; Lee et al., 2013) were carried out for three weeks while most of the others were four weeks long (Ertelt et al., 2007; Franceschini et al., 2010; Franceschini et al., 2012; Kim and Bang, 2016; Sale et al., 2014). The longest study was eight weeks long (Zhu et al., 2015).

In all the studies, except one (Zhu et al., 2015), the training took place five days a week.

Each session varied in length from a minimum of 10 min (Lee et al., 2013) to a maximum of 90 min (Ertelt et al., 2007); most of them were 30 min long (Franceschini et al., 2010; Franceschini et al., 2012; Kim and Kim, 2015a; Sale et al., 2014; Zhu et al., 2015). In two studies the session lasted 40 min (Franceschini et al., 2010; Kim and Bang, 2016) whereas in the last two the sessions they were 60 min long (Sugg et al., 2015) (60-90 minute long(Cowles et al., 2013)).

The length of the studies differed but all of them except one (Cowles et al., 2013) were less than an hour long.

Setting

The studies were held mainly in a laboratory or in an in-patient hospital (90,9 %), only one of the eleven articles selected was home-based (9,1%) (Sugg et al., 2015).

Outcome measures.

To investigate the body function domain of ICF, five studies used Fugg Meyer Assessment (FMA) (Franceschini et al., 2010; Kim and Bang, 2016; Sale et al., 2014; Sugg et al., 2015; Zhu et al., 2015) or modified Ashworth Scale (MAS) (Kim and Bang, 2016; Zhu et al., 2015). One other study used Motricity Index (MI) (Cowles et al., 2013).

All the studies focused on the ICF activity domain as shown in Fig. 3. In one study (Kim and Kim, 2015b) a kinematic analysis was carried out. The outcome measures used to investigate the activity domain were diverse; however, three studies used BBT (Franceschini et al., 2012; Kim and Bang, 2016; Sale et al., 2014), and three used Frenchay Arm Test (Cattaneo et al.; Ertelt et al., 2007; Franceschini et al., 2010; Franceschini et al., 2012). Three studies used the Barthel index (BI)(Franceschini et al., 2010; Kim and Bang, 2016; Zhu et al., 2015) (in the article by Kim and Bang;2016 and in the Franceschini et al; 2010 study the modified version was used) and two the Wolf Motor Function Test (WMFT)(Ertelt et al., 2007; Kim and Kim, 2015a). The other outcome measures can be seen in Table 5 and 6.

Only one study (Ertelt et al., 2007) analysed the participation domain using Stroke Impact Scale (SIS).

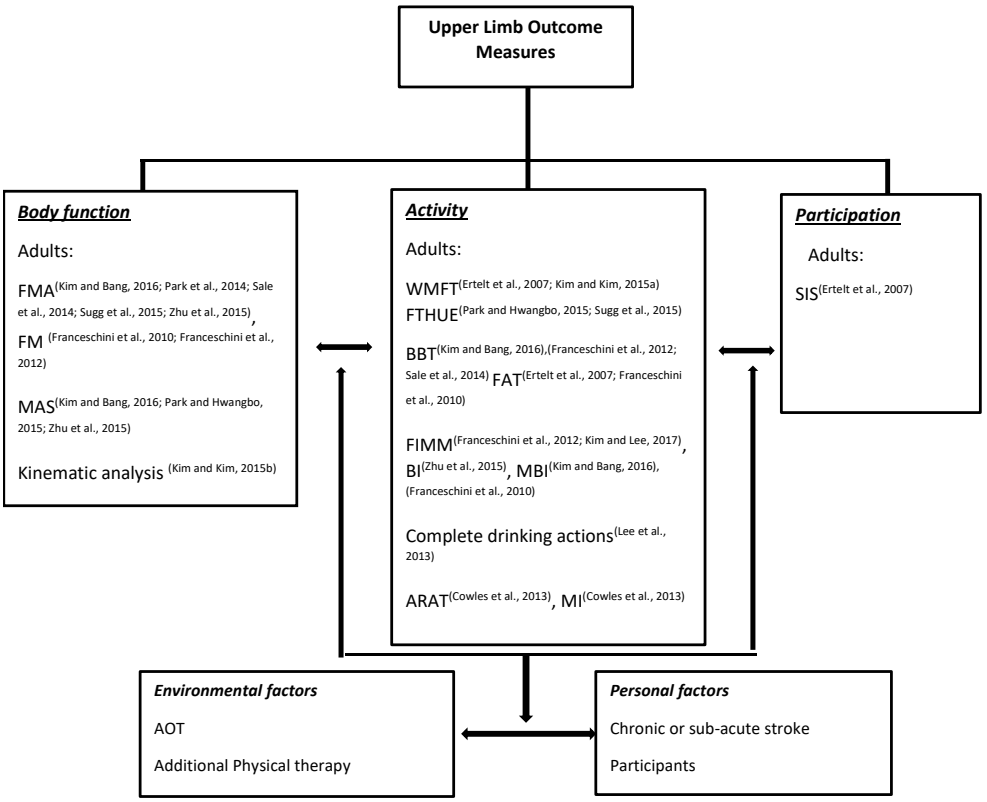


Fig.3

Effectiveness of AOT

Of the 11 studies carried out on adults on the upper limb eight found significant differences between groups at least in one outcome measure, in Franceschini et al. study (Franceschini et al., 2010) there was only one group and differences before and after treatment were observed; whereas two of the studies (Cowles et al., 2013; Kim and Kim, 2015b) were not significant. The patients enrolled differ in these two articles: in the study by Kim and Kim;2015 (Kim and Kim, 2015b) no information is available regarding patients whereas the patients in the study by Cowles et al; 2013 (Cowles et al., 2013) had had a stroke in the previous 31 days. The outcome measures used were kinematic analysis (average velocity and motor angle for the first) or Action Research Arm test

(activity ICF domain) or Motricity index investigating the body function domain. Structured interviews were held in a study (Sugg et al., 2015), and revealed that the participants found the AOT videos beneficial, and four participants reported the first instances of their movement becoming automatic since the stroke had occurred.

Statistic analysis on upper limb studies using ProMeta3

To analyse the overall effect size of the studies we analysed the outcome measures of all the studies regarding the upper limb, both on children and adults. The effect size was studied in a random-effect model rather than a fixed model because several differences were observed in the studies examined. Of the 15 articles selected regarding the upper limb it was not possible to insert two articles (Franceschini et al., 2010) (Kim and Kim, 2015b) because too few measures were reported. In two articles, where no means and standard deviations of the outcome measures were specified and those used were given in percentages, we converted them into decimal numbers to calculate the effect size. (Cowles et al., 2013; Sale et al., 2014). Thirteen different studies were analysed in three different projects according to ICF domains (i.e. body function, activity and participation). In the body function domain (Table 1 and Fig. 4) we analysed nine studies, with nine different outcome measures. The overall effect size was medium (0.49; CI 0.217- 0.766) and statistically significant ($p<0.001$). Twelve studies were analysed in the activity domain, with eleven different outcome measures (Table 2 and Fig. 5). The overall effect size was medium (0.52; CI 0.324 – 0.723) and statistically significant ($p < 0.001$). Only two studies investigated participation outcomes (Table 3 and Fig. 6). As these two studies were carried out on both adults and children it was not possible to combine them with a fixed model. The studies were therefore combined in a fixed-model resulting in a medium overall effect size (0.4; CI -0.308-1.106). However, it must be highlighted that the effect size of one of the studies was 0.01 and 0.8 of the other. The analysis performed was not statistically significant ($p < 0.315$).

Table 1. *ICF body function domain*

Author	Outcome measure	Effect size	Lower limit	Upper limit	Significance 2-tailed (p)	Variance	Standard error	Weight
Buccino G et al; 2012	MUUL	0.91	0.130	1.685	0.022	0.16	0.40	6.49%
Cowles T et al; 2013	Morticity Index	0.16	0.104	0.216	< 0.001	< 0.01	0.03	13.42%
Franceschini M et al; 2012	FM	0.13	-0.314	0.569	0.572	0.05	0.23	10.01%
Kim CH et al; 2016	FMA	1.35	0.428	2.281	0.004	0.22	0.47	5.33%
	MAS	0.62	-0.233	1.479	0.154	0.19	0.44	5.85%
Kim JY et al; 2014	Grasp power	0.36	-0.644	1.330	0.465	0.25	0.50	4.90%
Sale P et al; 2014	FMA	0.75	0.667	0.833	0.000	< 0.01	0.04	13.33%
Sgandurra G et al; 2013	MUUL	0.04	-0.763	0.837	0.928	0.17	0.41	6.30%
Sugg K, 2015	FMA	0.36	-0.390	1.103	0.349	0.15	0.38	6.76%
Zhu MH, 2015	FMA	0.33	-0.175	0.835	0.201	0.07	0.38	9.27%
	MAS Elbow flexional muscles	0.74	0.220	1.258	0.005	0.07	0.26	9.11%
	MAS Forearm pronator	0.53	0.016	1.037	0.043	0.07	0.26	9.20%
*Overall (random-effects model)		0.49	0.217	0.766	< 0.001	0.02	0.14	100.00%

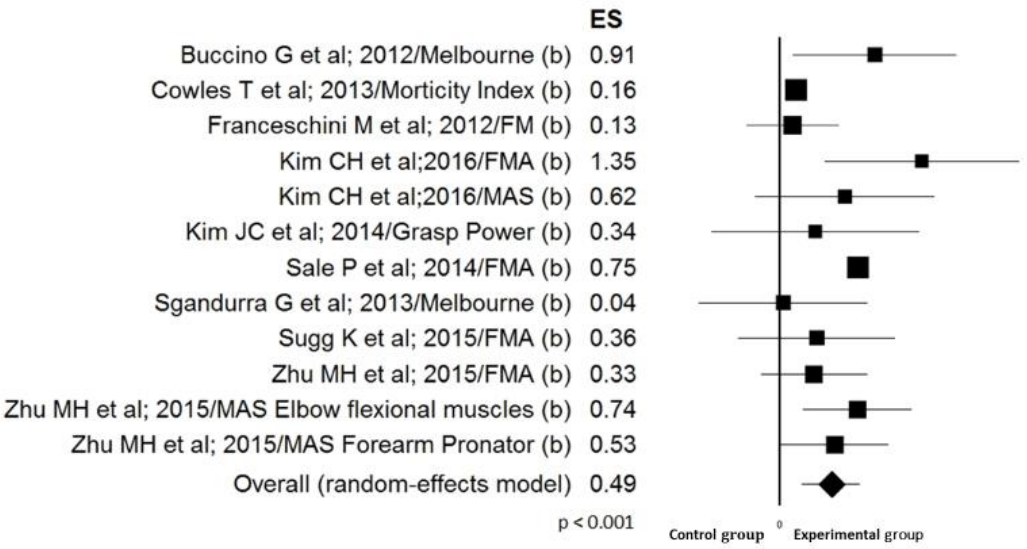


Figure 4 Icf body function domain

Table 2. *ICF activity domain*

Author	Outcome measure	Effect size	Lower limit	Upper limit	Significance 2-tailed (p)	Variance	Standard error	Weight (%)
Cowles T. et al; 2013	Action Research Arm Test	0.27	0.200	0.340	< 0.001	< 0.01	0.04	12.55%
Ertlet D. et al; 2007	FAT	2.38	1.097	3.667	< 0.001	0.43	0.65	2.04%
	WMFT activity	0.34	-0.650	1.324	0.35	0.25	0.50	3.10%
Franceschini M et al; 2012	BBT	0.30	-0,139	0.748	0.12	0.05	0.23	7.82%
	FAT	0.79	0,333	1.249	0.001	0,05	0,23	7.63%
	FIMM	0.13	-0.311	0.572	0.39	0.05	0.23	7.85%
Kim JY et al; 2014	ABILHAND kids	0.54	-0.455	1.541	0.20	0.26	0.51	3.05%
	BBT	0.53	-0.467	1.528	0.21	0.26	0.51	3.05%
Kim E et al;2015	WMFT	0.05	-0.750	0.851	0.63	0.17	0.41	4.18%
Kim CH et al; 2016	BBT	1.17	0.267	2.077	0.011	0.21	0.46	3.52%
	MBI	2.46	1.350	3.564	< 0.001	0.32	0.56	2.59%
Kirpatrick E et al; 2016	AHA	0.26	-0.260	0.775	0.23	0.07	0.26	6.87%
Lee D et al; 2013	number of drinking motion	0.62	-0.323	1.569	0.14	0.23	0.48	3.30%
Sale P et al; 2014	BBT	0.68	0.621	0.739	<0.001	< 0.01	0.03	12.60%
Sgandurra G et al; 2013	ABILHAND kids	0.23	-0.575	1.031	0.40	0.17	0.41	4.16%
	AHA	0.45	-0.364	1.256	0.19	0.17	0.41	4.11%
Sugg K et al; 2015	FTHUE	0.14	-0.599	0.884	0.49	0.14	0.38	4.62%
Zhu MH et al; 2015	BI	0.55	0.040	1.063	0.034	0.07	0.26	6.94%
*Overall (random-effects model)		0.52	0.324	0.723	<0.001	0.01	0.10	100.00

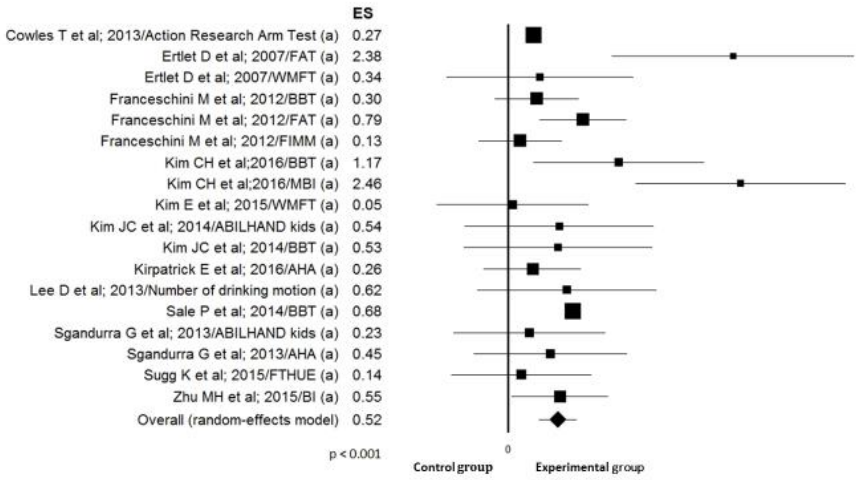


Figure 5 **Icf activity domain**

Table 3. *ICF participation domain*

Author	Outcome measure	Effect size	Lower limit	Upper limit	Significance 2-tailed (p)	Variance	Standard error	Weight
Ertlet D et al; 2007	SIS	0.82	-0.204	1.84	0.117	0.27	0.52	48.38%
Kim YC et al; 2014	WEE-FIM	0.01	-0.966	0.99	0.977	0.25	0.50	51.62%
*Overall (random-effects model)		0.40	-0.308	1.106	0.315	0.16	0.40	100.00%

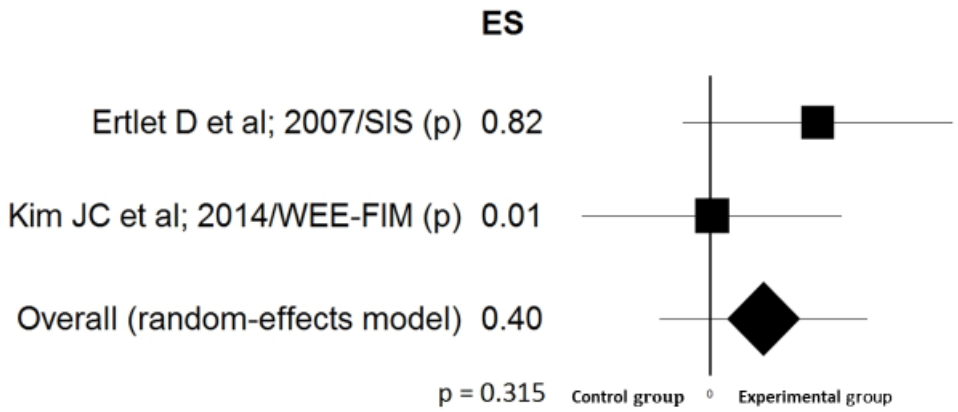


Figure 6 Icf participation domain

Studies focused on lower limb

Sample participants

All the eight studies on lower limb dysfunction were carried out on adult patients who had been diagnosed with chronic stroke.

In most of the studies the inclusion criteria were as follows:

1. Absence of visual or auditory problems, or aphasia (6/8) (Bae and Kim, 2017; Kim and Lee, 2017; Lee et al., 2017; Park and Hwangbo, 2015; Park et al., 2017; Park et al., 2014)
2. Ability to walk independently for 10 m. (6/8)
3. Mini-Mental State Examination score (MMSE) of >24 (6/8)(Kim and Lee, 2017; Kim et al., 2013; Kim and Lee, 2013; Park and Hwangbo, 2015; Park et al., 2017; Park et al., 2014)
4. No orthopaedic or cardiovascular disease; (5/8) (Kim and Lee, 2017; Lee et al., 2017; Park and Hwangbo, 2015; Park et al., 2017; Park et al., 2014)

The presence of neglect syndrome was an exclusion criterion (5/8)(Bae and Kim, 2017; Kim et al., 2013; Kim and Lee, 2013; Lee et al., 2017; Park et al., 2017)

In one study (Lee et al., 2017) a MMSE score of 23 or higher was included, while in another (Kim and Lee, 2017) patients were enrolled if they could independently stand up from sitting position and walk more than 30 metres with or without the use of an assistive device. In a further study (Bae and Kim, 2017), a grade 1 o 2 MAS score for the ankle and a poor to fair score in ankle dorsiflexion manual muscle testing (MMT) were also required as inclusion criteria.

AOT training and Control condition

Videos for the Lower Limb showed the action the participant was asked to perform (e.g. dorsiflexor exercises (Bae and Kim, 2017) or sit to walk exercises (Kim and Lee, 2017; Kim et al., 2013; Kim and Lee, 2013) or walk, which was either a simple movement carried out in a laboratory setting (Kim and Lee, 2013; Lee et al., 2017; Park and Hwangbo, 2015; Park et al., 2014) or an action in a particular environment (Park et al., 2017).

Generally, the models in the videos were healthy adult males and females but only in two articles (Kim et al., 2013; Kim and Lee, 2013) was their age specified (50 years) and this was similar to the mean age of the patients.

In order to minimize individual deviations, the video was produced separately for patients with left or right hemiplegia in two studies (Kim and Lee, 2013; Lee et al., 2017). The perspective was specified (Lee et al., 2017; Park et al., 2014) only in two studies and the speed of the sequence was reproduced in fast and slow motion in the front, back and side views in one (Park et al., 2014), in another study the action was presented at normal speed and half the normal speed (Park et al., 2017).

The content of the video-clips consisted in adults walking in different locations and grounds in two studies (Park and Hwangbo, 2015; Park et al., 2017), an

exercise of weight shift to the affected side was also included in another (Park et al., 2014).

In a study (Kim and Lee, 2017) the participants observed Sit To Walk(STW) video tasks and imitated the actions. The action observation tasks were composed of 16 tasks in which the difficulty and condition were adjusted based on the patient's functional status and level. In two studies (Kim et al., 2013; Kim and Lee, 2013) several of the stages in the video included trunk flexion, trunk rotation, sit to stand, and crossing obstacles for enhancement of balance and gait ability. In another (Lee et al., 2017) there were 3 stages of active assistive exercise: the first stage showed knee joint extensor and dorsiflexor training, the second stage knee joint flexor and dorsiflexor training and the third stage hip joint flexor.

In a further study (Bae and Kim, 2017), before of the application of electromyography-triggered functional electrical stimulation (ETFES) and the imitation of the movement, the dual afferent sensory input group (DASI) observed a 20 minute-video of dorsiflexion of the contralateral ankle recorded in advance. Simultaneously at the application of ETFES, movement of the contralateral ankle, induced by ETFES, was also shown live on a monitor while the subject was performing the action. Due to the DASI and the particular way in which the action observation training is conducted, it is very difficult comparing this study to the others and isolate the outcomes correlated to AOT.

The type of treatment given to the control group, where present, varied: five of them watched videos showing static landscapes (Kim and Lee, 2017; Park et al., 2017; Park et al., 2014)or pictures of nature (Park and Hwangbo, 2015). There were two control groups in three studies. In one of the control groups (MTA) (Lee et al., 2017) mirror therapy and the physical training of the same motions of AOT were carried out, while in the other control group action observation only was conducted, without any physical training.

In the other two studies (Kim et al., 2013; Kim and Lee, 2013), where two control groups were present, one group participated in a motor imagery

programme and did physical training based on contents similar to the AOT group, while the other had only done physical training.

In one study (Bae and Kim, 2017), AOT is combined with electromyography-triggered functional electric stimulation (ETFES) to improve voluntary functional movement and this is compared to the training of subjects in a control group who have undergone functional electric stimulation (FES).

Training length

Most of the interventions lasted 4 weeks (Bae and Kim, 2017; Kim et al., 2013; Kim and Lee, 2013; Park et al., 2017; Park et al., 2014). However, one study was 6 weeks long (Lee et al., 2017) and another was 8 weeks long (Park and Hwangbo, 2015).

In four studies the session lasted 30 min and took place three times a week (Kim and Lee, 2017; Lee et al., 2017; Park et al., 2017; Park et al., 2014). In one study (Kim and Lee, 2017), the 30 min session was divided into two of 15 min sessions/day (2 min + 30 sec for the observation phase plus 12 min and 30 sec for physical training).

In the other three studies the participants attended 30 min sessions, five times a week (Kim et al., 2013; Kim and Lee, 2013; Park and Hwangbo, 2015) whereas in one study the patients attended a 20 min session five times a week (Bae and Kim, 2017).

Setting

All the studies were conducted in a laboratory (Bae and Kim, 2017; Kim and Lee, 2017; Kim et al., 2013; Kim and Lee, 2013; Lee et al., 2017; Park and Hwangbo, 2015; Park et al., 2014) and the participants were recruited following examination of referrals from clinicians in an inpatient rehabilitation hospital (Lee et al., 2017; Park and Hwangbo, 2015; Park et al., 2017; Park et al., 2014). In one study(Park et al., 2017) the subjects watched each video-clip on a 23-inch computer monitor, which was positioned 150 cm in front of them. In a study (Lee et al., 2017), the AOT group sat comfortably in an armchair

and watched a video of motions performed by other people on the monitor installed 1 metre from them.

Outcome Measures

When assessing the lower limb, the main functions detected in most studies regarded the body function and the activity domain.

Balance was the main outcome for the body function domain. However, this was assessed differently in four studies (Bae and Kim, 2017; Kim and Lee, 2017; Lee et al., 2017; Park and Hwangbo, 2015) which made comparison of the studies and their relative outcomes difficult.

Measurements of gait were mainly used to explore the ICF activity domain such as: GAITRite system (Kim and Lee, 2013; Park et al., 2017; Park et al., 2014), 10 m walk test (Park and Hwangbo, 2015; Park et al., 2017; Park et al., 2014) and the figure of 8 walking test (Park et al., 2014). Gait ability was also investigated by analysing the results of questionnaires such as the Community walk test and the activities-specific balance confidence scale (Park et al., 2017), the Modified Functional Ambulation Profile (mE-FAP) (Lee et al., 2017) and the Walking Ability Questionnaire (WAQ) and Functional Ambulation Category (FAC) (Kim and Lee, 2013).

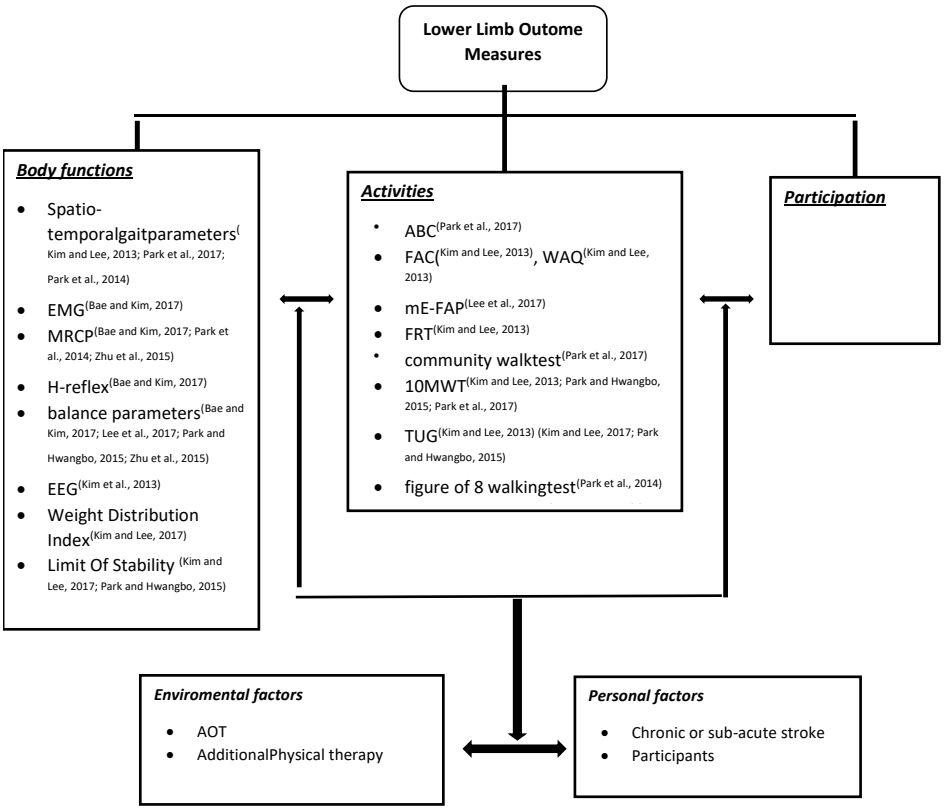


Fig.8

Effectiveness of AOT

Significant improvements in gait ability before and after AOT were found in five of the selected studies (Kim and Lee, 2017; Kim and Lee, 2013; Park and Hwangbo, 2015; Park et al., 2017; Park et al., 2014); Furthermore, three of them showed significant differences in more than one outcome measure, (Park and Hwangbo, 2015; Park et al., 2014) (Park et al., 2017).

AOT also improved balance in three studies (Kim and Lee, 2017; Lee et al., 2017; Park and Hwangbo, 2015) but no agreement was found on the variables of balance ability analysed and their outcomes.

In one study (Kim and Lee, 2017), which was a crossover randomized controlled trial in more than two groups, no significant differences in the Timed Up and Go Test (TUG), Dynamic Gait Index (DGI), and Weight Distribution

Index (WDI) were found between the Action Observation Physical Training (AOPT) and Landscape Imagery Observation Physical Training (LIOPT) groups ($p > .05$). However, the WDI, Limit Of Stability (LOS), DGI significantly increased during post-training 2, where the interventions were crossed over, as compared with post-training 1 in the LIOPT group ($p < .05$). These findings show that action observation could have an effect on the improvement of STW and balance abilities.

The significant improvement in the AOT group compared with the control group in patients' walking confidence indoors and outdoors (Park et al., 2017) and the significant changes in the two assessments of WAQ and FAC compared with the values before training (Kim and Lee, 2013) suggest that AOT can encourage social reintegration after stroke thanks to the anticipatory experience and preparation by observing a walking action.

In one particular article(Bae and Kim, 2017) changes in the cerebral cortex measured by the movement-related cortical potential (MRCP) showed improvement in both groups, but only in the DASI-group there was a significant difference in the motor potential (MP) at Cz and C4. The latter significantly differed in the DASI and the control group, together with the limit of stability (LOS) of dynamic balance. The results of measures such as H-reflex and muscle activity of tibialis anterior and medial gastrocnemius were not included because a greater interference of the ETFES intervention in our view was possible.

As regards the studies in which one of the two control groups was the motor imagery group, a significant difference was only found when comparing the action observation training group and the physical training group in TUG, gait speed, cadence and single limb support of the affected side (Kim and Lee, 2013), while in one study (Kim et al., 2013), no significant differences were observed in the three groups. However, in relative alpha power in Fp1 and Fp2 and relative beta power in Fp2 and C3, significant changes from EEG indicated that action observation training induced higher levels of cognitive activity than motor imagery or physical training. This can be explained if we consider that

patients after stroke may have difficulty forming clear representations of the imagined movement having not experienced it for a long time.
Due to the extreme differences in the outcome measures, it was not possible to examine the lower limb articles selected with the same procedure as those on upper limb function.

5.1.3 Discussion

AOT is a top-down rehabilitation approach whose aim is to promote the neuroplasticity and (re)-organization of the high-level brain circuits, such as the MNS, which are involved in motor behaviour. The main aim of AOT is, in fact, to directly induce a functional modification of the cortical components underpinning action organization, by exploiting the peculiar features of the MNS. The possibility of recruiting a crucial functionality of MNS, namely that of imitating by observation is among the major advantages of using this "top-down" model. To the best of our knowledge the pioneering studies on AOT in the field of neuro-rehabilitation were carried out in adult patients with stroke. Since the first AOT treatment was carried out by Ertelt in 2007 (Ertelt et al., 2007), the number of studies on adults and children has increased. Even if many reviews have been written (Bassolino et al., 2015; Buccino, 2014; Garrison et al., 2010; Johansson, 2011) to our knowledge, this is the first systematic review which also includes a meta-analysis, even if only for the upper limb, which systematically reports the results on the use of AOT in stroke patients. The overall grade of recommendation is B because most of the studies reviewed were level two, both for the upper and lower limb. An important finding of the present review is that preliminary evidence of the effectiveness of AOT on motor functions was found both in children with Unilateral Cerebral Palsy (UCP) and in adults and elderly people after stroke, from early phase trials. A further important and innovative method outlined in the present review, is the analysis of the effectiveness of AOT in relation to the

ICF domains. The general finding is that AOT significantly improves body function and activity domains, with significant medium effect size for the upper limb, in both children and adults.

As regards the setting in which the AOT is proposed, the present review found that AOT based on the observation of videos showing meaningful actions, followed by their execution, has promising results in many studies in patients at various ages from childhood to adulthood. Conversely, other studies have shown that observing a person rather than watching a video of AOT does not provide significant functional improvements either in adults with stroke in an early phase (mean 18.7 days) (Cowles et al., 2013) or in children with unilateral CP (Kirkpatrick et al., 2016). Therefore, it can be claimed that the observation of videos followed by the reproduction of the goal-directed action is the setting to be followed rather than observing a person performing the action. Furthermore, videos are not only easier to standardize but allow a broader range of patients to avail of therapy.

As regards the control group, if the aim of the study is to evaluate the effectiveness of AOT in terms of observation followed by motor replication rather than performing the action without watching a video, it follows that watching a video with no motor content is equivalent to performing the action without watching a video. However, it is difficult to compare the control group to those who do not perform any action at all. If, on the other hand, the aim of the study is to evaluate the effectiveness of AOT as a new rehabilitation method, no specific action is required of the control group.

Another important issue is the environment where the AOT can be delivered. All but two of the studies were carried out in hospital or in a laboratory setting (Kirkpatrick et al., 2016) (Sugg et al., 2015) which were delivered at home. These two home-based studies were carried out on different types of subjects (adults and children) and with different types of training (a video was used in one study (Sugg et al., 2015), while in the other a parent showed the action (Kirkpatrick et al., 2016)). In the study by Kirkpatrick et al. (Kirkpatrick et al., 2016) the authors explain the reasons for their choice as follows: “We chose

this model because of the straightforward nature of the approach and the potential to supplement therapy services, which currently cannot provide high-intensity upper limb input to all eligible children. However, monitoring of treatment fidelity is more challenging in this setting than in a clinic-based environment with therapist delivery". This issue is very relevant and well standardized home-based studies need to be encouraged as these would reduce not only hospital attendance and the waiting time for the therapy but would also allow a much greater number of patients to avail of the treatment. Taking this perspective into account, an ongoing trial on home-based AOT for upper limb in children with UCP, based on the use of technologies, is currently under way (www. Clinicaltrials.gov).

Future recommendations

The standardisation of AOT future research studies is fundamental as regards environment, treatment, control group and outcome measures.

AOT should be related to a more structured and targeted environment. Moreover, given the variability of the perspectives used in the videos, standardisation of these variables is needed in order to provide the most effective AOT. Furthermore, further studies comparing different settings (e.g. hospital versus home) are needed.

As regards the length and the treatment carried out when deciding the length of each teaching session, the question regarding the length of time a child can maintain his/her concentration should be posed. Furthermore, a time standardization is needed to understand if it relates to the AOT training outcomes. Finally, it would be useful to understand if there is a threshold to be reached before an effect is had on the MNS and if a total length is necessary to maintain the effect in the follow up.

As regards the outcome measure, a standardization, mainly for the lower limb and the evaluation of the ICF participation domain, is also recommended.

Finally, only a few studies in this review focus on children and none of those studies focused on the lower limb. It would be desirable if in the future a larger

number of studies investigated the lower limb function in children and identified a homogeneous population, e.g. the effect of AOT in a sample of patients suffering from bilateral CP, rather than investigate a larger sample including children suffering from Unilateral CP.

Limitations of this review

Only 23 articles matched the inclusion criteria of this systematic review of which only four studies had more than 50 patients enrolled. (Franceschini et al., 2012; Kirkpatrick et al., 2016; Sale et al., 2014; Zhu et al., 2015)

The studies selected used different types of inclusion and exclusion criteria resulting in very heterogenous populations across studies. Moreover, for the studies on children, different types of CP were included.

Due to the small and heterogeneous samples in the various studies it is difficult to draw conclusion on the results for chronic stroke patients. Moreover, as only four of the studies were carried out on children, results have to be interpreted with caution.

Hence, the evidence of AOT therapy is limited compared to other neurological conditions (e.g. Parkinson disease) (Caligiore et al., 2017). Finally, a follow up assessment is necessary to understand the long term changes after AOT.

Table 5a *Upper limb children*

AUTHOR	CEBM level	PEDro Score	Diagnosis	Sample size		Duration (wks)	Intensity (Frequency per wk; Minutes/day)
				Enrolled: tot;	Age (years)		
Kirkpatrick E et al; 2016	1b	6/10	UCP	70	3 to 10 years	12	5 sessions per week,15 minutes each session.
Kim JY et al; 2014	2b	5/10	CP (4 UCP 4 Bilateral CP)	16	9+/-3 years old	4	3 days a week, twelve 30-min sessions.
Sgandurra G et al; 2013	1b	8/10	UCP	24	5-15 years	3	15 consecutive working days, 60-min (including the rest periods) rehabilitation sessions.
Buccino G et al;2012	2b	7/10	Cerebral Palsy (12 UCP 3 bilateral CP)	15	6,8 years	3	5 times per week, 15 min/day;

AUTHOR	Experimental group				Control group		Outcome measures	Results: differences between experimental and control group	Findings
	Type of AOT	Type of actions	Videos (perspectives)	Other concurrent treatments/therapies	Type of intervention	Other concurrent treatments/therapies			
Kirkpatrick E et al; 2016	AOT: watch a parent perform a movement each time before attempting it.	/	Observation of a parent (sat next to the child, facing the same direction, and on the side of the less-affected hand) performing the movement each time before attempting it.	not specified	perform the purposeful action observation program without the observation phase.	not specified	Body function: MA2. Activity: AHA ABIL-HAND-Kids. Participation:NA	no between-group differences in AHA, MA2, or ABILHAND-Kids at 3 or 6 months vs baseline (all p>0.05)	Using a home-based parent-delivered play therapy approach, there is no difference in upper limb outcomes between AO+RP therapy and RP alone for children aged 3 to 10 years with UCP. However, low-intensity upper limb therapy can be delivered at home by parents, incorporated into play for as little as 1 hour per week with a small but sustained benefit
Kim JY et al; 2014	AOT: 12 possible videos, shown at normal speed, then twice at lower speed and then again at normal speed. The observation time for each video was five minutes (2,5min x2) and two videos were shown during each training session. After observing the videos, the subjects repeatedly practiced the same actions for 10 min per task.	Actions related to the children's daily lives	Three dimensions: front, sides and rear of the subject		Observation of photos of diverse landscapes such as mountains and the sea instead of actions. Then the same training as so for the AOT group.		Body Function : Grasp Power, Modified Ashworth Scale (MAS). Activity: BBT, ABILHAND-Kids; Participation: WeeFIM.	1) Grasp and MAS, which measure abilities the areas of physical functions and structures, showed significantly greater increases in the action observation physical training group than in the physical training group in the present study. 2) The two groups in the present study showed no significant differences in ABILHAND-Kids and WeeFIM.	A greater positive effects were seen on upper extremity functions in children with cerebral palsy in response to action observation physical training compared to simple physical training.
Sgandurra G et al; 2013	AOT: Observation of video sequences (3-min) of unimanual or bimanual goal-directed actions (each repeated at least 15 times) and subsequently executed observed actions with the hemiparetic UL or both ULs (3-min session) in the exact order and with the same objects as shown in the video.	15 sets of unimanual and bimanual daily life upper limb exercises each one composed by 3 sequential UL goal-related actions of increasing complexity.	First-person perspective	/	To watch computer games + verbally instructions to perform the same UL actions and in the same order as the experimental group.	/	Body function: MUUL; Activity: AHA,ABILHAND-Kids; Participation:NA	1) AHA= within-group differences were significant at all followup assessments (T1, T2, and T3) versus T0 (P = .005, .016, and .007, respectively). At the primary endpoint (T1) between-group difference in the changes was significant (P = .008) and this difference tended to remain at T2 and T3, although at the limits of statistical significance (T2 − T0, P = .019; T3 − T0, P = .049); 2) Melbourne and ABILHAND-Kids= no statistically significant differences.	1) Action observation training appears to be effective in improving affected UL daily actions with effects that persist at 6 months. 2) Our therapeutic protocol was inexpensive and time efficient and should be an easily reproducible rehabilitation program.
Buccino G et al;2012	AOT: To watch 9-12 min videos (everyday arm/hand actions), divided in 3-4 motor tasks, 3-min long each + to execute for 2 minutes what the patients had observed, to the best of their ability.	actions related to the children's daily lives	different perspectives	/	Video clips (no specific motor content), divided into three or four parts, each lasting 3 minutes + to execute the same actions as children in the case group for the same duration	/	Body function: Melbourne Assessment Scale. Activity:NA; Participation: NA.	At baseline groups did not differ on functional evaluation. After treatment, the functionalscore gain (D) was significantly different in the case and control groups (p=0.026)	1) AOT may play an important role in the rehabilitation of children with CP. 2) Not only did children in the case group show greater improvements than those in the control group on the Melbourne Assessment Scale, but this result was consistent in all participants.

AUTHOR	CEBM level	PEDro Score	Diagnosis	Sample size		Duration (wks)	Intensity (Frequency per wk; Minutes/day)	Outcome Measures	Results: differences between experimental and control group
				Enrolled: tot;	Age (years)				
Kim CH et al; 2016	2b	9/10	sub-acute stroke.	22	62,78 +/- 9,85 (AOT group); 61,49 +/- 8,64 (control group)	4	5 times/week, 40 min/day	Body function: FMA Activity: BBT, MBI, MAS, Participation:NA.	The mean change of FMA-UE, BBT, and MBI in the action observation training was significantly different between groups. The modified Ashworth scale (P>0.05; 95% CI, -0.402 to 0.624) did not show significantly different between groups.
Zhu M-H et al; 2015	2b	7/10	Stroke	61	42-75 years	8	Six times/week, 30 min/day,	Body function : FMA, MAS; Activity: BI; Participation:NA	FMA, BI and MAS scores were significantly better after treatment in the experimental group compared to the controls (all $p < 0.05$).
Kim E et al; 2015	2b	3/10	stroke.	12	Not specified	6	5 sessions per week, 30 min per day	Body function:NA Activity: WMFT; Participation:NA	The experimental group showed significantly greater improvement compared with the control group.
Kim E et al; 2015	2b	3/10	stroke.	12	Not specified	6	5 sessions per week, 30 min per session	Body function: Kinematic analysis ; Activity:NA, Participation:NA	1) The experimental group showed more improvement than the control group, but the difference was not statistically significant (p>0.05). 2) Both groups showed improvements in average velocity, trajectory ratio, and movement degree, but no statistically significant differences were observed between the groups.
Sugg K et al; 2015	4	/	Chronic stroke	15	68,86 ± 6,04	2	6 sessions (from the 18th to 28th day), 60-90 min each session.	Body function: FMA; Activity FTHUE; Participation:NA . 3) Confidence in Arm and Hand Movement Scale (CAHM scale)=; 4) Motor Activity Log (MAL); 5)Structured Interview Questions (improvements in function of UL , TRAINING ASPects participant thought contributed to functional improvements)	FMA, FTHUE(number of tasks): Significant difference in scores across time series testing points, and significant in the two phases but larger in the experimental group. FTHUE (level achieved): Significant difference in scores across time related to improvements only in the experimental group. FTHUE (Time average, seconds):not significant difference across times. CAHM: scores differed significantly across time and significant in the two phases. Structured Interview Questions: 13 participants reported that watching video demonstration of tasks was more beneficial while 1 felt as though they did not find a difference between the videos.

Table 6b: *Upper limb adults*

AUTHOR	Experimental group				Control group	
	Type of AOT	Type of actions	Videos (perspectives)	Other concurrent treatments/therapies	Type of intervention	Other concurrent treatments/therapies
Kim CH et al; 2016	Observation (9 minutes), followed by a break (1 minute to organize their thoughts, and practicing (30 minutes)	Task-oriented training consisted of performing task based on ADLs, e.g. folding up a towel, removing a bottle cap, lifting a cup and drinking from it.	Three views: front, side and top	Conventional rehabilitation program that involved occupational (1 hour/day), physical (2 hours/day), and speech therapies (as required)	Same tasks during a 30 minutes period, without watching the video.	conventional rehabilitation program that involved occupational (1 hour/day), physical (2 hours/day), and speech therapies (as required)
Zhu M-H et al; 2015	Observation followed by the action (total of 30 action videos used).	Many different movements e.g bending and extension, abduction and adduction, and pronation and supination of the shoulder joint; and more complex actions e.g. grabbing, catch and release of large and small balls, cubes and cylinders, holding and release of a coin and a key, putting on clothes.	Straight on (20s), right above(15 s) and right inside(15s); the compleate action was recorded 2-3 times at each angle.	Conventional drug treatment, traditional physical therapy and occupational therapy	Routine rehabilitation treatment and nursing	Conventional drug treatment, traditional physical therapy and occupational therapy
Kim E et al; 2015	observation of two activities per session selected by the patients themselves and repeated over 1 week.	Possible activities: feeding, grasping a can, combing hair, drinking water using from a cup, stamping, turning a page, mopping, folding towels, and opening a bottle cap using scissors)	Not specified	Occupational Therapy (not specified)	Perform the purposeful action observation program without the observing purposeful actions	Occupational Therapy (not specified)
Kim E et al; 2015	observation of two activities per session selected by the patients themselves and repeated over 1 week.	daily life activities	Not specified	Occupational Therapy (not specified)	Perform the AOT assignments, without the observational part	Occupational Therapy (not specified)
Sugg K et al; 2015	Observation of video clip (10 s, shown 3 times consecutively) followed by physical practice for 3 attempts. This process is repeated 10 times.. Five different motor tasks in each session.		Not specified	/	Relaxation-Sham Plus Physical Practice= observe video of neutral environmental scenery prior to the opportunity for physical practice of selected motor skills.	/

AUTHOR	CEBM level	PEDro Score	Diagnosis	Sample size		Duration (wks)	Intensity (Frequency per wk; Minutes/day)	Outcome Measures	Results: differences between experimental and control group
				Enrolled: tot;	Age (years)				
Sale P et al;2014	1b	8/10	Subacute ischaemic stroke	67	66,5 ± 12,7	4	5 days/week, two 15-min daily session at least 60-min interval apart	Body function: FMA; Activity: BBT; Participation: NA	The comparison between groups revealed a significantly higher gain for the EG than the CG, with respect to functional measures taken at both T1 and T2.. Moreover, Left hemiparetic (LH) subjects achieved significantly greater benefits compared to the right ones (RH). FMA: between groups, for RH no statistically significant differences at T1 and T2, for LH p=.003 atT1 and p=.029 for T2. BBT: between groups, for RH no statistically significant differences at T1 and T2, for LH p=.005 atT1 and p=.008 for T2.
Lee D et al; 2013	1b	5/10	Chronic stroke	33	Observation group (63 ± 3.7); Action Practice Group (62 ± 1.5); Combined group (61 ± 2.3); Control group (60 ± 5.9)	3	5 sessions per week, 10-min / day.	Body fnction:NA Activity: Complete drinking actions; Participation:NA	The action observation, combination, and action training groups showed statistically significant improvements compared to the control group ($p<0.05$). Combined group showed a significant higher number of drinking behaviours respect the Observation group ($p<.05$) immediately after the experiment and one week after. No statistical difference was found between the Combined and the Action training group.
Cowles T et al; 2013	2b	7/10	early after stroke	29	60-89 years	3	Each day for 15 working days, two 30-minute sessions (approximately 6- to 8-minute periods divided by 2 to 4 minutes of resting), separated by a 10-minute rest.	Body function:NA Acvitity: MI, ARAT; Participation:NA	The median (95% CI) between-group difference was not statistically significant.

Table 6d: *Upper limb adult*

AUTHOR	Experimental group				Control group	
	Type of AOT	Type of actions	Videos (perspectives)	Other concurrent treatments/therapies	Type of intervention	Other concurrent treatments/therapies
Sale P et al;2014	Observation followed by performing the same tasks (2 min). One task per day, starting from the easiest and ending with the most complex action).	20 daily activities e.g. drinking from a glass, combing hair, opening a box. Each task is composed by three different meaningful motor sequence displayed in order of ascending difficulty	First-person	Standard Rehabilitation: 3h/day (1 h physioterapy, 1 h occupational therapy, 1 h speech and swallow therapy).	Control Treatment: 5 static images displaying objects, without any animal or human being, for 3 minutes + to performe the same tasks of the experimental group	Standard Rehabilitation: 3h/day (1 h physioterapy, 1 h occupational therapy, 1 h speech and swallow therapy).
Lee D et al; 2013	1) Observation group: observation of a task video (20 times); 2) AOT group: observation of task video (5 minutes) followed by the actions (5 minutes)	Drinking behaviour: Action of stretching out the right hand to pick up a cup, bringing the cup to the mouth in order to touch the mouth, and then returning the cup to its initial position.	Front	/	1) Action practice group: repeatedly practiced the actions performed during the preliminary test for 10 minutes; 2) Control group: neither watched the video nor practiced the actions.	/
Cowles T et al; 2013	Observation (1 -2 minutes) followed by action (4 -6 minutes) performed simultaneously with the therapist who modulated timing as appropriate for the participant's ability			CPT as deemed appropriate, prior the intervention (Soft Tissue Mobilization, Joint Mobilization, Facilitate Muscle Activity, Positioning, Specific Sensory Input, Splinting Exercise to Increase Strength, Balance and Mobility, Upper Limb Functional Tasks, Education, Other)	CPT as deemed appropriate (Soft Tissue Mobilization, Joint Mobilization, Facilitate Muscle Activity, Positioning, Specific Sensory Input, Splinting Exercise to Increase Strength, Balance and Mobility, Upper Limb Functional Tasks, Education, Other). No other therapies	/

AUTHOR	CEBM level	PEDro Score	Diagnosis	Sample size		Duration (wks)	Intensity (Frequency per wk; Minutes/day)	Outcome Measures	Results: differences between experimental and control group
				Enrolled: tot;	Age (years)				
Francesch ini M. et al.;2012	1b	7/10	Stroke	102	Not specified	4	5 sessions/wk; 20 sessions (15 minutes =3-minute sequence observations and 2-minute action performances for 3 motor sequences); each session was repeated twice per day, at least 60 minutes apart. In between, the patient was requested to rest.	Body function:FM, MAS Activity: BBT, FAT,FIMM; Participation:NA.	Differences between the two froups were found from T0 to T1 and from T1 to T2. However, no difference was found on either change in BBT performance from T1 to T2. No significant difference between the study groups was found with regard to the FIMM and FM performance
Francesch ini M. et al; 2010	4	/	Chronic stroke	28	58,5 ± 14.0	4	5 days per week, each session lasted around 40 minutes.	Body function: FM, Activity: MBI,FAT (FIM scale.; Ashworth Scale (AS);) Participation:NA	/
Ertelt D. et al; 2007	1b	6/10	Stroke	16	38-69 years	4	18 consecutive working days, 18 sessions of 90 min each	Body function: NA; Activity:FAT, WMFT; Participation: SIS.	1) Significant improvement of motor functions in the course of a 4-week treatment, as compared to the stable pre-treatment baseline, and compared with a control grouphave been found. The improvement lasted for at least 8 weeks after the end of the intervention. 2) Neural activations between experimental and control groups after training yielded a significant rise in the bilateral ventral premotor cortex, bilateral superior temporal gyrus, the supplementary motor area.

Table 6f Upper limb adults

AUTHOR	Experimental group				Control group	
	Type of AOT	Type of actions	Videos (perspectives)	Other concurrent treatments/therapies	Type of intervention	Other concurrent treatments/therapies
Franceschini M. et al.;2012	Observation of 1 task per day consisting in three different 3-min motor sequences, starting from the easiest and ending with the most complex action) + to imitate the motor sequence they had observed.	20 daily activities e.g. drinking from a glass, combing hair, opening a box. Each task is composed by three different meaningful motor sequence displayed in order of ascending difficulty	First-person	1) Inpatient Rehabilitation= at least 3 h/d of physiotherapy	Control treatment or "sham" action observation= to observe for 5 min 5 static images (no motor contents) + to perform limb movements as well as feasible for 2 minutes according to a standard sequence, simulating those performed by the EG, involving shoulder and elbow joint mobilization.	1) Inpatient Rehabilitation= at least 3 h/d of physiotherapy
Franceschini M. et al; 2010	Observation (3 minutes) followed by performing the same tasks (2 min).	20 daily activities e. g To clean the table To take a clothes brush and brushing; To take a soap and wash the hands; To take and eat an apple; To take and to change a position of a jar	Different perspectives	/	/	
Ertelt D. et al; 2007	Observation (6 minutes) followed by action (6 minutes) of the observed actions with their paretic upper limb using the same objects as shown in the video film. Every day a 'unit' of three hand and/or arm movements of increasing complexity were presented.	daily routine activities	Three different persepectives.	/	The same of AOT group with the exception that the patients watched sequences of geometric symbols and letters instead of action sequences. The practiced hand and arm actions were performed by instruction of the assisting therapist in the exact order as they were practiced under the experimental condition.	/

Table 7a. Lower limb

AUTHOR	CEBM level	PEDro Score	Diagnosis	Sample size: Enrolled: tot: experimental group(s); control group(s)	Sample size: Age (years)	Duration (wks)	Intensity (Frequency per wk; Minutes/day)	Outcome Measures	Results: differences between experimental and control group
Kim J-C et al; 2017	1b	8/10	Chronic stroke	21 (11 AOPT + 10 LIOPT)	AOPT : 57.08 +/- 7.29; LIOPT: 52.92 +/- 8.21	3 ws + 3ws = 6	3 days/wk.,15 min x 2 /day; AOPT: (2 min 30 sec for video + 12 min + 30 sec for physical training) x 2/day	Body function: WDI, LOS ; Activity: TUG , Dynamic Gait Index Participation: NA	No significant difference in the TUG, DGI, and WDI were found between the AOPT and LIOPT groups ($p > .05$). There was a significant difference in LOS between the AOPT and LIOPT groups ($p < .05$).
Bae S et al; 2017	1b	7/10	Crhonic stroke	18 (9 DASI group) + 9 (Ctrl only FES)	DASI : 49.50 +/- 10.60 ; Ctrl : 49.67 +/- 8.78	4	5 d/ws, 20 min/d.	Body function: Balance, H-reflex, EMG, MRCP Activity:NA; Participation:NA	MRCP in MP at C4 and dynamic balance (LOS) showed significant differences between DASI and control group ($p<0,05$)
Park HJ et al; 2017	1b	7/10	Chronic stroke	25 (12 Exp gr + 13 Ctrl gr)	Exp gr (57.33 +/- 6.89) ; Ctrl gr (55.08 +/- 8.12)	4	3 sessions per week, 30 min for video,	Body function:spatiotemporal gait measure. Activity: ABC(activities-specific balance confidence scale), 10MWT, community walk test Participation:NA.	In the experimental group, walking function and ambulation confidence was significantly different between the pre- and post-intervention, whereas the the control group showed a significant difference only in the 10-m walk test
H. J. Lee et al; 2017	2b	5/10	Chronic stroke	35	AOTA gr. (Exp gr): 62.8 +/- 7.4 ; Ctrl gr (MTA gr. :57.27 +/- 5.7 ; AOT gr.: 59.8 +/- 6.7)	6	3 times per week , 30 minutes/day.	Body function: Balance; Activity: mEFAP; Participation:NA	No significant difference was found between the groups ($p>0.05$) on OBI, ABI, MBI measures or in the fall risk and about the mEFAP.

AUTHOR	Experimental group				Control group		Findings
	Type of AOT	Type of action	Videos (perpesctives)	Other Concurrent Treatments	Type of intervention	Other Concurrent Treatments	
Kim J-C et al; 2017	Observation of tasks related to STW and imitated the actions. There were 16 tasks in which the difficulty and condition were adjusted based on patient's functional status and level.	The video tasks consisted of "Wheel without back of chair, or Handle without wheel of chair, walk 3 m, return, and sit in different surface	Not specified		Observation of static landscape photos, such as mountains, beaches, valleys, waterfalls, and countryside. After that, a physical training (the same of AOPT) was given.		The WDI, LOS, DGI significantly increased during posttraining 2, in which the interventions were crossed over, as compared with posttraining 1 in the LIOPT group ($p < .05$). These trainings showed that the action observation could have an effect on the improvement of STW and balance abilities.
Bae S et al; 2017	Observation of a 20 minute-video of dorsiflexion of the contralateral ankle recorded in advance; while simultaneously at the application of ETFES, movement of the contralateral ankle, induced by ETFES, was also shown live on a monitor during subjects' performance	Ankle Dorsiflexion	Not specified	A Microstim device was used to apply FES and in addition an EMG-electrode was placed on the TA muscle to detect muscle activity (ETFES). General physical therapy including Bobath approach for 30 min/d, 5 d/wk was also given.	Patients were instructed to dorsiflex upon FES application. A Microstim device was used to apply FES by bipolar placement of the electrodes. Asymmetrical biphasic waves were applied for 20 min with valgus position.	General physical therapy including Bobath approach for 30 min/d, 5 d/ws	DASI stimulated voluntary movement in patients, caused rapid activation of the cerebral cortex, and reduced excessive excitation of SMN.
Park HJ et al; 2017	The subjects watched each video clips on a 23-inch computer monitor, positioned 150 cm in front of them	Walking on even and uneven ground,in a complex and unpredictable community environment, in a parking lot, shopping centre.	two different filming speeds: in normal speed and half times the normal speed.	Functional training based on sit-to-stand balance, therapist-guided walking training and stair climbing for 30 min, 5 times per week, for 4 weeks.	30 min video clips of static landskape pictures; any human or animal representation were excluded	Same Functional training based on a sit-to-stand balance training, therapist-guided walking training and stair climbing for 30 min, 5 times per week, for 4 weeks	AOT may be useful for improving walking skills outdoors and indoors. AOT can facilitate gait- associated self-efficacy.
H. J. Lee et al; 2017	Dorsiflexor training composed of three stages of active assistive exercise in which they watched video for 15 minutes a day through the monitor installed 1 m away from them and executed the same physical training for 15 minutes a day.	1 stage: knee joint extensor and dorsiflexor training. 2 stage: knee joint flexor and dorsiflexor training. 3 stage: hip and knee joint flexor and dorsiflexor training	the front and lateral side videos were produced separately for the left and right hemiplegic subjects.	General physical therapy 2 times/wk for 30 minutes/day	The MTA group received mirror therapy for 15 minutes a day and physical training of the same motions without a mirror for 15 minutes a day. The AOT group conducted action observation only, without physical training, for 30 minutes a day.	general physical therapy 2 times/wk for 30 minutes/day	AOTA improved subjects' static balance and gait function because of the OBI and ABI significantly decreased, such as mEFAP, while MBI and the fall risk did not significantly differ but decreased.

Table 7c: *Lower limb*

AUTHOR	CEBM level	PEDro Score	Diagnosis	Sample size: Enrolled: tot: experimental group(s); control group(s)	Sample size: Age (years)	Duration (wks)	Intensity (Frequency per wk; Minutes/day)	Outcome Measures	Results: differences between experimental and control group
Park et al; 2015	2b	4/10	Chronic stroke	40 (20 AOGT + 20 GGT)	AOGT: 51.15 +/- 14.81 ; GGT : 48.65 +/- 12.81	8	5 times per week, 30 minutes per session.	Body function: Balance,limit of stability Activity: TUG, 10MWT. Participation:NA	There were significant differences in the sway speed, in the limit of stability, in TUG and 10 MWT between the two groups after the experiment ($p<0.05$), but not in the sway area ($p>0.05$).
Park HR et al; 2014	2b	7/10	chronic stroke	21 (n=11 to the experimental group + n=10 to the control group)	Exp gr: 55.91 +/- 9.1 ; Ctrl gr 54.80 +/- 12.22	4	3 times per week, 30-min/day.	Body function: Spatiotemporalgait paramenters; Activity :10MWT, Figure-of-8 walk test ; Participation:NA	The difference between the pre- and post-test values of the 10-m walk test, figure-of-8 walk test, and dynamic gait index were statistically significantly different between the groups ($p < 0.05$).
Kim JH et al; 2013	2b	6/10	Chronic stroke	27 (n=9 to the AOT group; n=9 to the motor imagery group; n=9 to the physical training group)	AOT gr: 55.3 +/- 12.1; MIT gr 54.8 +/- 8.8 ; PT gr : 59.8 +/- 8.9	4	5 times / wk, 30-min traininig for session.	Body function: EEG; Activity:NA; Participation:NA.	there were no significantly differences between the three groups.
Kim JH et al; 2013	2b	6/10	Chronic stroke	27 (n=9 to the AOT group; n=9 to the Motor Imagery group; n=9 to the Physical Training group)	AOT gr: 55.3 +/- 12.1; MIT gr 54.8 +/- 8.8 ; PT gr : 59.8 +/- 8.9	4	5 times / wk, 30-minute training for session.	Body function:Spatiotemporal gait parameters Activity: TUG, the functional reaching test(FRT). Participation:NA, the walking ability questionnaire(valuta la loro mobilità dando 5 punti: non riesco, assistito...), and the functional ambulation category (gait ability)	No significant differences in any of the outcome measures were observed between the action observation training group and the motor imagery training group, except for Stride length. While significant difference was observed between the action observation training group and the physical training group in the TUG, gait speed, cadence, and single limb support of the affected side

Table 7d: *Lower limb*

AUTHOR	Experimental group				Control group		Findings
	Type of AOT	Type of action	Videos (perpesctives)	Other Concurrent Treatments	Type of intervention	Other Concurrent Treatments	
Park et al; 2015	AOGT(Action Observation Gait Training): 3-min videos each for different walking with 1-min break; + 5-min walking training each on the same flat land, slope, and steps as in the video with 1–2 minute break.	Walking on a flat land, on a slope, and on steps	Not specified	30 minutes of general physical therapy, prior to each training (joint exercise, muscle strengthening and stretching exercise.)	GGT (General Gait Training): 12-min video with break (3 min) showing images of nature unrelated with walking + 20-min walking training.	30 minutes of general physical therapy, prior to each training (joint exercise, muscle strengthening and stretching exercise.)	Action observational training had a positive effect on static balance and gait ability.
Park HR et al; 2014	Observation of 4 tasks for functional walking for 10min and after 20-min sessions of walking training.	The 4 tasks including weight shifting to the affected side, walking on straight and curved paths, walking on even and uneven surfaces, crossing obstacles, and walking with functional tasks	2 filming speed options (normal speed and 2-times slowerthan-normal speed) in the front, back and side views in twice sequence.	/	Observation of video clips showing different landscape images for 10 mi followed by the performance od the same walking tasks of the AOT group	/	AOT can be beneficial in enhancing walking ability of patients with post-stroke hemiparesis, and may be clinically feasible as a practical adjunct to routine rehabilitation therapy.
Kim JH et al; 2013	They watched a task video for 20 minutes, followed by physical training with a therapist for 10 minutes based on the video	4 stages including trunk flexion, trunk rotation, sit to stand, and crossing obstacles	Not specified	Neurodevelopmental therapy for 30 minutes, twice per day, 5 days per week for 4 weeks	Motor Imagery Group: 20 minutes of motor imagery program played through a computer speaker + physical training for 10 minutes based on the training contents. Physical Training Group: training of the trunk for learning supine to rolling movements, sit to stand, and normal gait pattern.	Neurodevelopmental Therapy (only in the Motor imagery group)= 30 minutes, twice per day, 5 days per week, 4 weeks	In alpha power, a significant decrease in Fp1 and in Fp2, was observed in the action observation group, as in beta power, a significant decrease in Fp2 and in C3. AOT induced higher levels of cognitive activity than motor imagery or physical training did.
Kim JH et al; 2013	They observed a task video for 20 minutes through a 32-inch TV installed approximately 2m away while sitting in a comfortable armchair, followed by physical training with a therapist for 10 minutes, based on the video.	Stage 1) pelvic tilting, trunk flexion and extension, and trunk rotation in the sitting position; Stage 2) sit to stand and stand to sit; Stage 3) weight shift to the front and back, left and right; Stage 4) gait level surface and step over obstacle	the video was produced separately for patients with left hemiplegia and those with right hemiplegia	Neurodevelopmental therapy for 30 minutes, twice per day, 5 days per week for 4 weeks	Motor Imagery Group: 20 minutes of motor imagery program played through a computer speaker + physical training for 10 minutes, identical to the contents suggested in the action observation training program. Physical Training group: training of the trunk for learning supine to rolling movements, sit-to-stand, and normal gait pattern, as well as training of thelower extremity, weight shifting, and gait level surface and gait stairs	Neurodevelopmental Therapy (only in the Motor imagery group)= 30 minutes, twice per day, five days per week	Although no significant difference was observed between the groups, the action observation training group showed the numerically largest changes, compared to the other groups. Action observation training and motor imagery training results in improvement of dynamic balance and gait ability.

5.1.4 References of the review

- Bae, S., Kim, K.Y., 2017. Dual-afferent sensory input training for voluntary movement after stroke: A pilot randomized controlled study. NeuroRehabilitation 40, 293-300.
- Bassolino, M., Sandini, G., Pozzo, T., 2015. Activating the motor system through action observation: is this an efficient approach in adults and children? Developmental medicine and child neurology 57 Suppl 2, 42-45.
- Bellelli, G., Buccino, G., Bernardini, B., Padovani, A., Trabucchi, M., 2010. Action observation treatment improves recovery of postsurgical orthopedic patients: evidence for a top-down effect? Arch Phys Med Rehabil 91, 1489-1494.
- Brunner, I.C., Skouen, J.S., Ersland, L., Gruner, R., 2014. Plasticity and response to action observation: a longitudinal FMRI study of potential mirror neurons in patients with subacute stroke. Neurorehabilitation and neural repair 28, 874-884.
- Buccino, G., 2014. Action observation treatment: a novel tool in neurorehabilitation. Philosophical transactions of the Royal Society of London. Series B, Biological sciences 369, 20130185.
- Buccino, G., Arisi, D., Gough, P., Aprile, D., Ferri, C., Serotti, L., Tiberti, A., Fazzi, E., 2012. Improving upper limb motor functions through action observation treatment: a pilot study in children with cerebral palsy. Developmental medicine and child neurology 54, 822-828.
- Buccino, G., Binkofski, F., Fink, G.R., Fadiga, L., Fogassi, L., Gallese, V., Seitz, R.J., Zilles, K., Rizzolatti, G., Freund, H.J., 2001. Action observation activates premotor and parietal areas in a somatotopic manner: an fMRI study. The European journal of neuroscience 13, 400-404.

- Buccino, G., Vogt, S., Ritzl, A., Fink, G.R., Zilles, K., Freund, H.J., Rizzolatti, G., 2004. Neural circuits underlying imitation learning of hand actions: an event-related fMRI study. Neuron 42, 323-334.
- Caffarra, P.e.a., 2016. The effectiveness of action observation treatment (aot) in alzheimer's disease: benefit on temporal orientation and visuo-prassic abilities. Alzheimer's & dementia: the journal of the alzheimer's association 12, 615.
- Caligiore, d., mustile, m., spalletta, g., baldassarre, g., 2017. Action observation and motor imagery for rehabilitation in parkinson's disease: a systematic review and an integrative hypothesis. Neurosci biobehav rev 72, 210-222.
- Cattaneo l, c.f., jezzini a, rizzolatti g., 2009. Representation of goal and movements without overt motor behavior in the human motor cortex: a transcranial magnetic stimulation study. The journal of neuroscience : the official journal of the society for neuroscience.
- Cattaneo, l., fasanelli, m., andreatta, o., bonifati, d.m., barchiesi, g., caruana, f., 2012. Your actions in my cerebellum: subclinical deficits in action observation in patients with unilateral chronic cerebellar stroke. Cerebellum (london, england) 11, 264-271.
- Cebm, oxford centre for evidence-based medicine – levels of evidence (march 2009).
- Cohen, j., 1977. Statistical power analysis for the behavioral sciences. Academic press, new york.
- Cowles, t., clark, a., mares, k., peryer, g., stuck, r., pomeroy, v., 2013. Observation-to-imitate plus practice could add little to physical therapy benefits within 31 days of stroke: translational randomized controlled trial. Neurorehabilitation and neural repair 27, 173-182.
- Ertelt, d., small, s., solodkin, a., dettmers, c., mcnamara, a., binkofski, f., buccino, g., 2007. Action observation has a positive impact on rehabilitation of motor deficits after stroke. Neuroimage 36 suppl 2, t164-173.

- Fadiga, l., fogassi, l., pavesi, g., rizzolatti, g., 1995. Motor facilitation during action observation: a magnetic stimulation study. Journal of neurophysiology 73, 2608-2611.
- Franceschini, m., agosti, m., cantagallo, a., sale, p., mancuso, m., buccino, g., 2010. Mirror neurons: action observation treatment as a tool in stroke rehabilitation. European journal of physical and rehabilitation medicine 46, 517-523.
- Franceschini, m., ceravolo, m.g., agosti, m., cavallini, p., bonassi, s., dall'armi, v., massucci, m., schifini, f., sale, p., 2012. Clinical relevance of action observation in upper-limb stroke rehabilitation: a possible role in recovery of functional dexterity. A randomized clinical trial. Neurorehabilitation and neural repair 26, 456-462.
- Garrison, k.a., winstein, c.j., aziz-zadeh, l., 2010. The mirror neuron system: a neural substrate for methods in stroke rehabilitation. Neurorehabilitation and neural repair 24, 404-412.
- Gatti, r., tettamanti, a., gough, p.m., riboldi, e., marinoni, l., buccino, g.,, 2013. Action observation versus motor imagery in learning a complex motor task: a short review of literature and a kinematics study. Neurosci lett. 540, , 37–42.
- Hoare, b., imms, c., randall, m., carey, l., 2011. Linking cerebral palsy upper limb measures to the international classification of functioning, disability and health. Journal of rehabilitation medicine 43, 987-996.
- Johansson, b.b., 2011. Current trends in stroke rehabilitation. A review with focus on brain plasticity. Acta neurologica scandinavica 123, 147-159.
- Kim, c.h., bang, d.h., 2016. Action observation training enhanced upper extremity function in sub acute stroke survivor with moderate impairment: a double-blind, randomized controlled pilot trial, journal of the korean society of physical medicine, pp. 133-140.

- Kim, e., kim, k., 2015a. Effect of purposeful action observation on upper extremity function in stroke patients. Journal of physical therapy science 27, 2867-2869.
- Kim, e., kim, k., 2015b. Effects of purposeful action observation on kinematic patterns of upper extremity in individuals with hemiplegia. Journal of physical therapy science 27, 1809-1811.
- Kim, j.c., lee, h.m., 2017. The effect of action observation training on balance and sit to walk in chronic stroke: a crossover randomized controlled trial. Journal of motor behavior, 1-8.
- Kim, j.h., chung, e.j., lee, b.h., 2013. A study of analysis of the brain wave with respected to action observation and motor imagery: a pilot randomized controlled trial. Journal of physical therapy science 25, 779-782.
- Kim, j.h., lee, b.h., 2013. Action observation training for functional activities after stroke: a pilot randomized controlled trial. Neurorehabilitation 33, 565-574.
- Kim, j.y., kim, j.m., ko, e.y., 2014. The effect of the action observation physical training on the upper extremity function in children with cerebral palsy. Journal of exercise rehabilitation 10, 176-183.
- Kirkpatrick, e., pearse, j., james, p., basu, a., 2016. Effect of parent-delivered action observation therapy on upper limb function in unilateral cerebral palsy: a randomized controlled trial. Developmental medicine and child neurology 58, 1049-1056.
- Lee, d., roh, h., park, j., lee, s., han, s., 2013. Drinking behavior training for stroke patients using action observation and practice of upper limb function. Journal of physical therapy science 25, 611-614.
- Lee, h.j., kim, y.m., lee, d.k., 2017. The effects of action observation training and mirror therapy on gait and balance in stroke patients. Journal of physical therapy science 29, 523-526.
- Mulder, t., 2007. Motor imagery and action observation: cognitive tools for rehabilitation. Journal of neural transmission 114, 1265-1278.

- Park, e.c., hwangbo, g., 2015. The effects of action observation gait training on the static balance and walking ability of stroke patients. Journal of physical therapy science 27, 341-344.
- Park, h.j., oh, d.w., choi, j.d., kim, j.m., kim, s.y., cha, y.j., jeon, s.j., 2017. Action observation training of community ambulation for improving walking ability of patients with post-stroke hemiparesis: a randomized controlled pilot trial. Clinical rehabilitation 31, 1078-1086.
- Park, h.r., kim, j.m., lee, m.k., oh, d.w., 2014. Clinical feasibility of action observation training for walking function of patients with post-stroke hemiparesis: a randomized controlled trial. Clinical rehabilitation 28, 794-803.
- Pomeroy, v.m., clark, c.a., miller, j.s., baron, j.c., markus, h.s., tallis, r.c., 2005. The potential for utilizing the "mirror neurone system" to enhance recovery of the severely affected upper limb early after stroke: a review and hypothesis. Neurorehabilitation and neural repair 19, 4-13.
- Rizzolatti, g., fadiga, l., gallese, v., fogassi, l., 1996a. Premotor cortex and the recognition of motor actions. Brain research. Cognitive brain research 3, 131-141.
- Rizzolatti, g., fadiga, l., matelli, m., bettinardi, v., paulesu, e., perani, d., fazio, f., 1996b. Localization of grasp representations in humans by pet: 1. Observation versus execution. Experimental brain research 111, 246-252.
- Rizzolatti, g., sinigaglia, c., 2016. The mirror mechanism: a basic principle of brain function. Nature reviews. Neuroscience 17, 757-765.
- Sale, p., ceravolo, m.g., franceschini, m., 2014. Action observation therapy in the subacute phase promotes dexterity recovery in right-hemisphere stroke patients. Biomed research international 2014, 457538.

- Sgandurra, g., ferrari, a., cossu, g., guzzetta, a., fogassi, l., cioni, g., 2013. Randomized trial of observation and execution of upper extremity actions versus action alone in children with unilateral cerebral palsy. Neurorehabilitation and neural repair 27, 808-815.
- Sugg, k., muller, s., winstein, c., hathorn, d., dempsey, a., 2015. Does action observation training with immediate physical practice improve hemiparetic upper-limb function in chronic stroke? Neurorehabilitation and neural repair 29, 807-817.
- Thieme, h., morkisch, n., rietz, c., dohle, c., borgetto, b., 2016. The efficacy of movement representation techniques for treatment of limb pain--a systematic review and meta-analysis. The journal of pain : official journal of the american pain society 17, 167-180.
- Zhu, m.h., wang, j., gu, x.d., shi, m.f., zeng, m., wang, c.y., chen, q.y., fu, j.m., 2015. Effect of action observation therapy on daily activities and motor recovery in stroke patients. International journal of nursing sciences 2, 279-282.

5.2 Tele-UPCAT: Study Protocol of a Randomized Controlled Trial of a Home-Based Tele-monitored UPper Limb Children Action Observation Training for Participants with Unilateral Cerebral Palsy

A new rehabilitative approach, called UPper Limb Children Action Observation Training (UP-CAT), based on the principles of Action Observation Training (AOT), has provided promising results for Upper Limb rehabilitation in children with UCP. This study will investigate if a new Information and Communication Technology (ICT) platform, named Tele-UPCAT, is able to deliver AOT in a home setting and will test its efficacy on children and young people with UCP.

The trial has been approved by the Tuscany Paediatric Ethics Committee (169/2016). Publication of all outcomes will be in peer-reviewed journals and conference presentations. Trial registration: ClinicalTrials.gov: NCT03094455 (16 March 2017). The trial was funded by Italian Ministry of Health grant: GR-2011-02350053

5.2.1 Methods

Study design

The Tele-UPCAT trial is an exploratory randomized, allocation concealed (waitlist-controlled), and evaluator-blinded clinical trial with two investigative arms using an AOT intensive rehabilitation program of home based AOT compared to standard care in children and young people with UCP. The study is a waitlist controlled trial, in order to allow all enrolled participants to perform AOT training either immediately or after a waitlist period. After obtaining informed consent, and completing baseline assessment (T0) participants will be block randomized into pairs according to the House functional classification system (HFCS) activity level (grades 2-3, 4-5 and 6-

8) and age (5-14y, 15-20y), (House et al, 1981) using a computer-generated set of random numbers. Randomization, sequence generation and preparation of group allocation materials will be carried out by an independent researcher who will be not involved in the trial. Pairs will be divided randomly into two groups with 1:1 experimental/standard care (waitlist) ratio. Participants allocated in the experimental group will immediately start AOT for a 3-weeks period, while those in the standard care group will continue with their usual care.

In both cases, AOT or standard care, participants and parents will be asked to take daily notes on a predefined diary of their daily activities, including therapies for their motor disability. In addition, to record the acceptability and feasibility of the training, participants and/or families allocated in the experimental group will fill in a multiple-choice questionnaire (with box for notes) for ascertaining the perception about their experiences of using the Tele-UPCAT system.

All participants will be re-evaluated after the period of experimental training/standard care (T1) with standardized tests and questionnaires (see outcome measures). T1 will be the primary endpoint aimed at evaluating the short-term effects of AOT according to CONSORT Guidelines (see Figure 1) (Cobos-Carbo and Augustovski, 2011).

After this phase, participants previously allocated to the experimental group will continue standard care, while who's who started with standard care will then commence home based AOT. The participants of this SC group will be re-assessed at the end of training (T1 plus). Further assessment of all participants will be performed after 8 weeks (T2) and 24 weeks (T3) from the end of AOT training (T1 or T1 plus, for the experimental and waitlist group, respectively), to evaluate the medium and long term effects of AOT. All the assessments will be carried out at home by trained therapists.

In summary, assessments will be performed at:

- T0, baseline: the week before the period of AOT/Standard Care
- T1: at 1 week after period of AOT/Standard Care
- T1 plus: the week after period of AOT, for waitlist group

- T2: 8 weeks after end of AOT
- T3: 24 weeks after end of AOT.

The details of the study design are reported according to CONSORT guidelines (Figure 1), SPIRIT (Standard Protocol Items: Recommendations for Intervention Trials) statement (Chan et al, 2013) and TIDier (Template for Intervention Description and Replication) Checklist (Hoffmann et al, 2014). The programme of enrolment, interventions and assessments designed according to SPIRIT guidelines are shown in Figure 2.

Blinding

All the clinical outcome measures (AHA, Melbourne Assessment 2 MA-2 and Box & Block Test) will be videotaped by a therapist blind to group assignment. Videotapes will be randomized and scored by assessors blind to group allocation and order of assessments. During each assessment all the participants will wear two Actigraphs (wGT3X-BT, wActiSleep-BT), one for each wrist.

Two independent researchers (two child neurologists) without competing interests will comprise the data monitoring committee for this study. They will review all adverse events (deciding to stop the trial if necessary), subject participant retention in each study arm and compliance with study protocol at 12 weekly intervals. Participants will have a study number in a dedicated data file. The file with study participants numbers and personal data will be stored in a password protected file, accessible only by the principal investigator. In order to promote participant's retention, all the assessment will be completed at home. The clinical primary and secondary outcome measures will be completed within one week of when expected. Enrolled participants are ethically able to withdraw from the study at any time, and have been notified their usual follow-up and clinical care would not be impacted.

Study sample and recruitment

Enrolment and clinical trial management will be carried out by child neurologists and physiatrists at the Department of Developmental Neuroscience of IRCCS Fondazione Stella Maris (FSM, Pisa, Italy), with the collaboration of the Unit of Children Rehabilitation of S.Maria Nuova Hospital (Reggio Emilia, Italy).

Potential participants will be identified according to inclusion criteria (see below), from UCP patients of the clinical departments. Suitable participants and their parents will be invited to participate and will be enrolled in the RCT only after written consent has been obtained.

Inclusion criteria are participants with:

- confirmed diagnosis of spastic motor type UCP;
- aged between 5 and 20 years at time of recruitment;
- predominant UL spasticity;
- House functional classification system, (HFCS) score ≥ 2 that is, able to passively hold an object in the hand or better
- cognitive level within normal limits i.e. Intelligence Quotient ≥ 70, as assessed in the last year prior to recruitment on the WPPSI-III, WISC-IV or WAIS
- participants and parents willing to commit to the intensive therapy program for a 3-week period.

Participants will be excluded in case of:

- previous orthopaedic surgery or Botulinum Toxin A (BoNT-A) injection in the UL within 6 months prior to study entry.

Sample size

Even if planned as an exploratory study, a sample size estimates, according to CONSORT guidelines, have been based on projected treatment effect on the primary outcome measure, AHA. Taking account of the study design and the stratification, a minimum sample size of 10 per group will be required in order to detect a 1.40 effect size (value based on our preliminary data) at significant

level of 0.05 and 80% power. Considering a 20% of possible drop-outs a minimum of 12 participants per group will be recruited, total sample of 24 participants.

Study treatment

The Tele-UPCAT system has been designed through the close collaboration between the rehabilitation staff (child neurologists and child therapists) of IRCCS Fondazione Stella Maris and biomedical engineers of BioRobotics Institute Scuola Superiore Sant'Anna. Taking into account the previous clinical experience on UPCAT, the main components of the Tele-UPCAT system (e.g. the size of the screen, the need of a guide for alternating the time of observation and of execution, the key words for catching the attention, etc), the AOT library of exercises (e.g. the adapation of the objects for enlarging the exercises to more impaired hands) and the experimental training (e.g. time and duration) have been defined. In general, the training will be structured in one session per day, to be executed 5 working days for 3 consecutive weeks (i.e. 15 sessions in total). The duration of daily sessions will be about 60 minutes per day for a total of 15 hours. Participants undergoing the AOT intervention through the Tele-UPCAT system will watch 3 minutes first-person video sequences of unimanual or bimanual goal-directed actions followed by their execution for 3 minutes. Each day 3 different actions will be proposed twice.

Tele-UPCAT system

Tele-UPCAT system (see Figure 3) has been designed and built and will be provided at home by the BioRobotics Institute of Scuola Superiore Sant'Anna (Pontedera, Italy).

It is a dedicated platform for delivering AOT designed to be user-friendly, by subjects at home in a playful setting with integrated smart features.

The platform has been designed and developed by integrating two different modules:

- The Observation Module (OM) for the presentation of AOT videos and recording of participant's attention and exercise execution. This consists of a computer with 23" desktop, a dedicated software and a video camera. The Observation Module has been obtained by integrating a large all-in-one personal computer (All-in-One touch HP EliteOne 800 G2 - L3N93AV), a large switch and a video camera (GoPro HERO Session), which will record a whole field, including subject's face and hands and table with objects. The Observation Module is important to determine whether the participant is looking at the monitor during observation phase and has an overall view of the execution of actions. A dedicated software, designed after a deep and specific literature analysis, was developed for guiding and motivating participants through the phases of AOT (observation followed by execution). In addition, the software was customized for the wide age range of participants providing an interactive game with an engaging story different for every day of training for school aged children, and a slide-show with a voice-over for adolescents and young people. The general architecture of the software is based on the following sequence: observation of a 3-minute video followed by execution of the same action for 3 minutes. Subsequently, the same video will be replayed and then executed a second time. As stated before, a 60-minute session, including rest intervals, of three different goal-directed actions of increasing complexity are observed and imitated twice every day. At the end of each day the software will terminate the session and automatically update it for the next day.
- The Motor Performance Module (MPM) for the execution of actions. This will be mainly composed of a kit of exactly the same common objects and toys shown in the videos and two Actigraphs (wGT3X-BT, wActiSleep-BT, for more details see http://actigraphcorp.com/support/activity-monitors/gt3x/) worn one for each wrist. With this design it will be possible to measure the upper limb activity during the AOT training while the lack of sensorized toys embedded in the MPM will not allow to measure quantitative measures of hand activity (e.g. force or pressure) during the AOT training.

The first prototype of Tele-UPCAT system has been widely tested before the beginning of the RCT in order to test the stability and reliability of the system.

AOT library

On the basis of the previous AOT exercises, (Sgandurra et al, 2011 and 2013) rehabilitation staff (child neurologist and child therapists) has created a library of rehabilitation packages composed of three different series of AOT exercises suitable to be executed at home. They differ for complexity of action and range of UL capabilities conceived in relation to HFCS levels (≤ 4, 5-6, 7-8). Each series is organized into customized sequences designed to cover unimanual and bimanual UL goal-directed actions with a variety of objects and toys commonly used in routine life. For each series, experimental training is composed of 15 sets (8 unimanual followed by 7 bimanual) of routine UL activities, to be completed in 3 weeks (5 days per week). Each set has a general common goal (e.g. drinking a glass of water) composed of three sequential tasks of increasing complexity (e.g. from picking up a bottle of water, to opening the cap, to pouring the water into a glass). As previously indicated, in order to grade the activities according to the range of capabilities, three series of sets have been planned. The actions of each series have the same goal but the material and type of movement (i.e., range of movement, type of grasp) is customized in order to guarantee feasibility of the proposed activity while maintaining the same overall objective (see Table 1 and Figure 4). Each action of the three series performed by an actor is videotaped so that the videos show only the hand and arm from the first perspective; each video is then edited to last 3 minutes. A right-handed actor uses one or two hands for unimanual and bimanual exercises respectively for participants with right UCP. For the left UCP the previous videos where reversed if they maintained the same characteristic of the setting and of the hand movement, while the remaining videos where specifically videotaped.

Table 1. List of goal-directed actions planned for the AOT training grouped in unimanual (white cells) and bimanual (grey cells) actions

Days	Action a	Action b	Action c
1	Uncover a little ball by lifting a box	Place a little ball in a glass	Fill a glass with water
2	Pick coloured card and match it to the same colour	Move a coloured card and place it on a base	Pick a card and place it on the similar one
3	Pick up a rubber stamp and move from/to different positions	Pick up a rubber stamp and press it against horizontal plane to print a figure	Pick up a rubber stamp and press it against sloping plane to print a figure
4	Pick up coin, put it in piggy bank through the slot on the top OR Pick up a magnet and place it on a horizontal magnetic board	Pick up coin, put it in piggy bank through the vertical slot on the side OR Pick up a magnet and place it on a sloping magnetic board	Pick up coin, put it in piggy bank through the horizontal slot on the side OR Pick up a magnet and place it on a vertical magnetic board
5	Pick up a wooden rubber stamp and move to different positions	Pick up a wooden rubber stamp and press it against horizontal plane to print a figure	Pick up a wooden rubber stamp and press it against sloping plane to print a figure
6	Move a spray can OR Move the bottle with a little ball inside	Place the spray can on a support OR Remove a little ball from the bottle	Put the spray can into a cup OR Press the catapult and launch a little ball
7	Move a container filled with shimmy powder	Open the container	Sprinkle shimmery powder on a paper
8	Place magnetic fish on a paper	Pick up fishing rod and catch magnetic fish	Pick up magnetic fish and place them in a container
9	Move a hole punch	Insert a sheet of paper and make holes	Match holes on sticks
10	Wet a cloth placing it in a container with water	Wring cloth and place it in a plate	Open a toy washing machine and insert the cloth inside

11	Pick up a card and place it on a support	Pick up a card and insert it in a clothespin	Pick up a card and insert it in a clothespin in a different orientation
12	Pick up and handle a piece of Play-Doh	Divide it in two pieces	Open a toy oven and insert a saucepan (with Play Doh in it)
13	Search for coin in the bag and place it on a support	Take the coin and insert it in a wallet	Open a box and place the wallet inside
14	Open a tube of tempera paint	Wet a brush with tempera paint	Make figure using a stencil with the brush dipped in tempera paint
15	Move a glitter glue tube	Open it	Decorate a frame by pasting pieces of mosaic

Experimental training

Before delivering the Tele-UPCAT platform, the training will be customized individually for each participant. The rehabilitation staff will select, on the basis of age and HFCS level, from the library the most appropriated AOT rehabilitation packages for each participant, then the engineers will upload them in the Observational Module (OM). For the Motor Performance Module (MPM), the therapists will organize a container of all the objects identifying them with numbers relative to the training day (e.g. little ball number 1 which means day 1 of the training). In addition, a dedicated printed manual with instructions and guidelines related to the different steps of the training and for system management and the setup will be provided. The manual contains also all the contacts of both technical and rehabilitation staff for remote assistance in case of any problems during the training. Two Actigraphs (wGT3X-BT, wActiSleep-BT) will be initialized for the recording period (3 weeks) to be worn on each wrist.

The ICT platform will be delivered to the participant's home by the engineers that are in charge of the installation of the system. The families will identify a designated position with a table or a desk of about 80x100 cm near to a socket where the ICT platform will be placed. Engineers and rehabilitation staff will train both parents and participants about the correct use of the system,

including safety aspects. During the first two training days, a therapist will visit each participant and their parents to confirm the set-up.

During the training sessions, each participant will sit on a chair with both arms placed on a table in front of a platform positioned at about 1 m. Especially when the participant will be a child, a parent will be seated on her/his more impaired side to prompt attention during task execution and assist if necessary. The software will guide the participant in the sequence of observations and executions.

Standard care

Participants previously allocated in the standard care group will continue their usual care for 3 weeks. Usual care for recruited participants could be consisted for physical or occupational therapy. The frequency and the type of all therapies will be recorded accurately by a diary in both groups.

Outcome measures

Description of sample

Children participating in the study will be classified according to HFCS, which assesses function of the impaired hand in children with CP. This classification consists of 9 grades ranging from a hand that is not used at all (grade 0) to one that is used spontaneously and independently from the other one (grade 8). Due to the general approach in classifying hand functional level, this scale can also be easily applied to young adults with UCP. HFCS will be used for all ages as a criterion for inclusion in this study (from grade 2 to grade 8). In addition, they will be classified according to the Manual Ability Classification System (MACS), a classification system of the child's ability to handle objects in daily activities on one of five levels.

Primary outcome measure

On the basis of our scientific hypothesis and according to previous clinical experience, the primary outcome measure will be the AHA. The latest version

5.0 will be used. This assessment measures UL function during bimanual activities by evaluating spontaneous use of assisting hand during a semi-structured age-appropriated 10-15-minute session with specific toys or objects requiring bimanual handling. The school-kids form will be used for the assessment of UCP children 6-12 years old while the Adolescent version (Ad-AHA), using the board game "Go with the Floe", will be completed with participants older than 13 years. This last version, even if validated up to 18 years, will be used with potential participants 18-20 years old to guarantee the same assessment across all participants regardless of age. Moreover, AHA has been already used, even if not validated, in young people with UCP (46, 47). The scale uses a Rasch measurement model which is a method to convert raw scores into a linear measure located on a unidimensional scale and more specifically to convert them into 0 to 100 logit-based AHA units, that will be used for the statistical analyses. All AHA assessments will be videotaped in a standardised manner and the subsequent scoring will be carried out by a certified expert rater who will be masked to group allocation and assessment order.

Other outcome measures

Other secondary measures will include measures of unimanual capacity (MA-2, and BBT) and bimanual daily activities at home and in the community (ABILHAND-Kids). Moreover, participation and quality of life will also be assessed. All assessment will be performed at T0, T1, T1 plus, T2 and T3 unless otherwise indicated. Questionnaires will be completed by parents and/or participants at home and if doubts will occur, child neurologists or therapists will be available to discuss face to face items not clear to them.

i. The Melbourne Assessment 2 (MA2) measures unilateral UL function and it is a valid and reliable tool for evaluating quality of UL movement in children with neurological conditions for ages between 2.5 and 15 years. MA2 is a criterion-referenced test that extends and refines the scale properties of the original Melbourne Assessment (MUUL) and like MUUL it can also be used

for adolescents and young adults. MA2 measures four elements of UL movement quality: movement range, accuracy, dexterity and fluency. It comprises 14 test items of reaching, grasping, releasing and manipulating simple objects. The test is administered by video recording the child's performance for subsequent scoring (30 items score). A raw score is provided for each of the four sections (Movement Range, Accuracy, Dexterity and Fluency) that will be analysed separately. It predominantly includes concepts within the body function domain as well as in the activity domain. Even if the MUUL and also the MA-2 have been validated up to 15 years, the first one has been used in studies involving patients with CP older than 15 years and in adults. (46-47) We have chosen to use the MA-2 also for participants older than 15 years to guarantee the same assessment across all participants regardless of age instead of using other scales (e.g. the Fugl-Meyer Assessment or the Action Research Arm test).

ii. Box and Block Test (BBT) is a quick (2-5 minutes), simple and inexpensive test which measures unimanual dexterity in the activity domain. BBT is composed of a test box with a partition in the middle and 150 wooden blocks (25mm). The patient had to transport as many blocks as possible in 1 minute from one compartment to another. Firstly, the patient is asked to perform the test with the unaffected hand and then with the affected hand. The number of blocks transported by affected hand in 1 minute will be counted and considered for the main analyses. It can be used for a wide range of populations from childhood to adulthood.

iii. ABILHAND-Kids is a semi-structured item-response questionnaire on a 3-point ordinal scale (impossible, difficult, easy) that measures daily manual activities referred to in the activity domain of ICF. Parents will be instructed to rate their child's perceived difficulty in performing each activity taking account the performance of their child when performing the activity without technical or human assistance, regardless of the limb(s) and the strategies used. It has been validated for children aged 6-15 years but it has been used for larger ranges (6-19 years). The questionnaire has been developed using

the Rasch measurement model which provides a method to convert the ordinal raw scores into a linear logit measures located on a unidimensional scale, that will be used for the analyses. This questionnaire will be used at all assessment periods.

iv. Participation and Environment Measure - Children and Youth (PEM-CY). It is a parent-reported instrument that evaluates participation and environment across home (ten items), school (five items) and community (ten items) settings. For each item, the parent is asked to identify how frequently (over the past four months) the child has participated (eight options: daily to never); how involved the child typically is while participating (five points scale: very involved to minimally involved); and whether the parent would like to see the child's participation in this type of activity change (no or yes, with five options for the type of change desired). For each setting, the parent is then asked to report on whether certain features of the environment make it easier or harder for the child to participate. The following summary scores will be obtained: total score and score for each of the three setting-specific environmental supportiveness (home, school, community). Moreover, the total number of supports and the total number of barriers will be computed. This questionnaire will be used at T0 and T3.

v. Cerebral Palsy Quality of Life Questionnaire for Children (CP QOL –Child, 4-12 years) and Cerebral Palsy Quality of Life Questionnaire for Adolescents (CP QOL –Teen, 13-18 years) evaluate quality of life in children and adolescents with CP. A score on a 0-100 scale will be obtained for each of computed sub-domains. In particular the Children form filled in by the parents assesses 7 subdomains (Social well-being and acceptance, Feelings about functioning, Participation and physical health, Emotional well-being and self-esteem, Access to services, Pain and impact of disability and Family health) and five subdomains (excluding Access to services and Family health to the previous) for the children version. The Teen form evaluates 7 subdomains (General well-being and participation, Communication and physical health, School well-being, Social well-being, Access to services,

Family health and Feeling about functioning) for the form filled in by the parents and 5 (excluding Access to services and Family health to the previous) for those filled in by the caregivers. These questionnaires will be used at T0 and T3.

vi. Quantitative measurement during unimanual and bimanual manipulation tasks, executed after the observation of the same tasks, is a new assessment tool that consists of observation, followed by execution of three tasks of increasing difficulty (unimanual lifting, bimanual placing near and bimanual cooperation, holding and pulling) by means of a sensorized object. New technological tools such as sensorized objects can help in assessing the manipulation capabilities (reaching and grasping) in a quantitative but ecological way and the sensitivity to a training. Previous studies of the authors were focused on the development of sensozied toys for measuring infant's manipulation (50-52). Starting from this experience, a new sensorized object has been designed and developed by the engineers tuning the sensors sensitivity and working range to the needs of participants with UCP.

Two load-cells and a switch embedded in the sensorized object allow for the measurement of the following parameters: grasping time, maximum grasping force and delay time between unaffected and affected hand in reaching for the object. This set-up is out of Tele-UPCAT system even if it was designed and developed in parallel.

vii. Quantitative measurement of bimanual activities will be performed in all the participants enrolled in the study by means of Actigraph GXT3+, as components of Motor Performance Module of Tele-UPCAT system. Actigraphs wGT3X-BT and wActiSleep-BT, equipped with a Velcro strap bracelet, will be worn, one for each wrists. As general rule, the experimental group have to wear the actigraphs not only during the training sessions but also between T0 and T1 and between T1 and T2 (total 6 weeks, 24 hours per day or as much as possible) while the control group are requested to wear them between T0 and T1, between T1 and T1 plus and, if possible also

between T1 plus and T2 (total 9 weeks, 24 hours per day or as much as possible). All the daily activities, experimental training or usual care, or removal will be recorded in a dedicated diary. The Actigraphs will also be worn, by all participants, during each time point of clinical assessments with the outcome measures. Data will be mainly relative to the asymmetry index (AI), that is the difference between the mean activity of the dominant with those of the non dominant hand and it will be correlated with the clinical scores obtained in the clinical outcome measures (mainly AHA).

viii. Cost effectiveness: A within trial cost-utility analysis will be conducted to synthesise the costs and benefits of the training program. Resource use (staff time, equipment and facility use) associated with the program will be collected alongside the RCT. Health care utilisation will be collected using a resource use questionnaire previously used in CP studies. Health utility will be derived from the adapted CHU-9D, a quality of life measure designed specifically for economic evaluation and which has been validated in an Australian population.

ix. A semi-guided face to face interview about the acceptability of the training will be completed immediately after the training period (T1 or T1 plus). It will be performed by the rehabilitation staff with the aim of investigating participant's and parents' opinions about the training in terms of customization of exercises, suitability to UCP children, feasibility at home, required effort by the participants and acceptability of Actigraphs, suitability of the manual and of the software will be recorded.

Statistical analyses

Clinical data will be analysed by means of the Statistical Package for Social Sciences (SPSS). Means and standard deviation of clinical outcome scores for both groups will be calculated to identify potential baseline differences between groups. As a first step, normality of distributions will be verified by Shapiro-Wilk's test. Between-group differences for all selected outcome measures will be evaluated at T0, by means of t-test for unrelated samples or

non-parametric Mann–Whitney U independent sample test, for normal or non-normal distributed data, respectively. To test our first hypothesis, an intention to treat analysis will be carried out, evaluating between-group differences (delta scores) for primary and secondary outcome measures at the primary endpoint (T1), compared with T0 (T1-T0), by means of parametric or non-parametric tests for unrelated samples. The age, HFCS level, characteristics of usual care (in both groups), supervision of caregiver, acceptability of therapy (measured by semi structured interview) will be analyzed for further secondary exploratory analyses in order to determine if some of these variables are predictive of better responses to the Tele-UPCAT training. In addition, a matched-pairs test (t-tests or Wilcoxon) will be carried out in order to assess retention of effects at follow-up periods (T1 or T1 plus, T2 and T3) relative to assessment before AOT training (T0 or T1 for experimental or waitlist group respectively). Bonferroni corrections will not be carried out in relation to the exploratory nature of the current RCT study and the relative small number of the study sample. To detect if significant changes will correlate to HFCS levels, a correlation analysis between score changes after AOT training (T1 or T1 plus) and assessment before AOT training (T0 or T1) will be carried out. Finally, an exploratory within-group analysis will be performed for the waitlist group comparing changes during AOT with respect to those of the first standard care period.

5.3 The tele-upcat project RCT: a pilot study on a sample group of children with unilateral cerebral palsy

Cerebral palsy is a major cause of physical disability and hemiplegia is the most frequent form. This disability represents a serious challenge for the patients' families both from an economic and clinical point of view. Unfortunately, not many therapies are available for the rehabilitation of the upper limb, which is important for a variety of daily activities such as grasping. Following the discovery of the mirror neuron system (MNS), AOT, an innovative type of rehabilitation treatment based on MNS plasticity has been proposed. This treatment has two phases: an action observation phase followed by the action performance. The positive impact of AOT on stroke patients is highlighted in literature but further studies which focus on children are needed. Moreover, of the studies carried out using this type of treatment only two were carried out at home and besides the one that focused on children was not significant.

The pressing need to conduct treatment at home comes from the observation that only those patients living near the hospital can usually attend an intensive period of treatment of three weeks.

When a child is undergoing intensive treatment he /she needs to be seen and monitored on a daily basis. This impacts greatly on the family and the parents who very often have to change their lifestyle which, in some instances, may mean moving for a period of time close to the hospital at a cost, taking holidays or time off work.

Thanks to the use of some technological tools e.g. a computer with a specific software and Actigraphs, patients were afforded the possibility of following the treatment at home and it was possible to broaden the population that could benefit from the AOT.

5.3.1 Methods

The aim of this pilot study was to verify the validity and feasibility of a protocol based on a new ICT platform, named Tele-UPCAT (Upper Limb Children Action Observation Training), to provide the action observation therapy in children with Upper Limb (UL) impairment at home, before starting the major randomized control trial. Moreover, it aims to measure the efficacy of the treatment in a sample group of children and adolescents with Unilateral Cerebral Palsy (UCP), comparing some preliminary data before and after treatment. Data was further compared to the data of other previous studies carried out in a clinical setting to evaluate whether treatment at home was more advantageous than treatment in a hospital.

Study design

This pilot study was carried out before the Tele-UPCAT project RCT. Six patients who met the inclusion criteria of the RCT study were selected. Moreover, the six patients we selected who had different grades of House Functional Classification System (HFCS), two for each grade (2-3), (4-5), (6-7). The patients were evaluated after three weeks of treatment T1, thus evaluating the short term effect of AOT.

The Tele-UPCAT RCT trial, which is currently underway, is an exploratory randomized, allocation concealed (waitlist-controlled), and evaluator blinded clinical trial. The study compares the AOT home-based intensive rehabilitation programme to standard care in children and adolescents with UCP. Participants are enrolled, according to the inclusion and exclusion criteria, and the informed consent is obtained. After the selection, they do a baseline assessment (T0) and then they are blocked randomized into pairs, using a computer-generated set of random numbers, according to the HFCS activity level (grades 2-3, 4-5 and 6-8) and age (5-14yrs, 15-20yrs).

Two groups of pairs have been randomly formed with 1:1 experimental/standard care (waitlist) ratio. The participants who have been

allocated to the experimental group start AOT for a 3 week period immediately, while those in the standard care group continue with their usual care.

Participants and parents in both groups are asked to keep note of their daily activities, in a predefined diary, including therapies for their motor disability. In addition, the participants and/or families in the experimental group are asked to fill in a multiple choice questionnaire (with a box for notes) to grade their feelings regarding the proposed activities so as to record the acceptability and feasibility of the training.

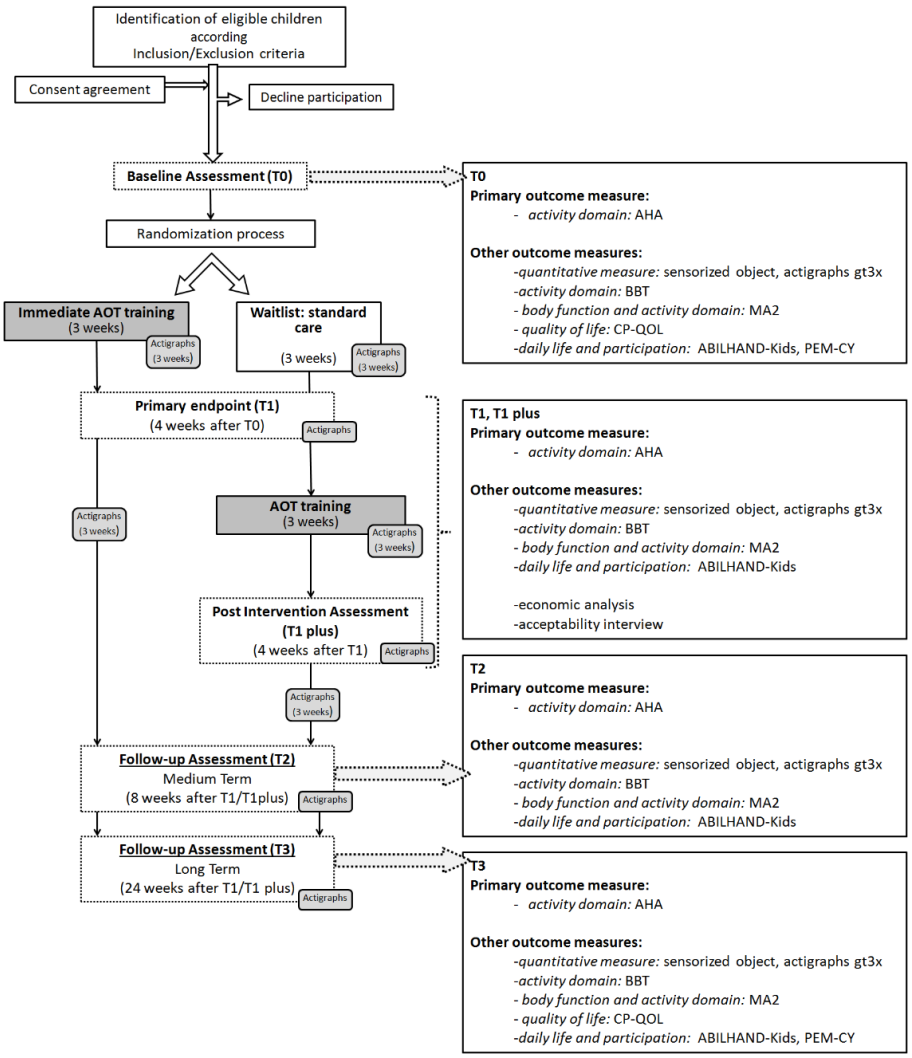


Fig.10 *RCT Study design*

The patients are evaluated after three weeks of treatment T1, thus evaluating the short term effect of AOT. In the RCT trial, the participants, who had been previously allocated to the experimental group, continue with standard care after this initial three week period of AOT, while those who started with standard care commence home-based AOT. The participants of this SC group are re-assessed at the end of training (T1 plus). A further assessment of all participants will be performed in the RCT after 8 weeks (T2) and 24 weeks (T3) from the end of AOT training (T1 or T1 plus, for the experimental and waitlist group, respectively), to evaluate the medium and long term effects of AOT. All the assessments will be carried out at home by trained therapists.

Study sample

Participants were selected from a database of patients with UCP by child neurologists and physiatrists at the Department of Developmental Neuroscience of IRCCS Fondazione Stella Maris (FSM, Pisa, Italy), with the collaboration of the Unit of Children Rehabilitation of S.Maria Nuova Hospital (Reggio Emilia, Italy), were obliged to meet the inclusion and exclusion criteria. The inclusion criteria were: (1) patients with a diagnosis of spastic hemiplegia, (2) the subject age between 5 and 20 years at T0 (3) predominant upper limb spasticity (4) HFCS within a level from 2-8 (5) cognitive level within normal limits i.e. Intelligence Quotient ≥ 70, as assessed in the last year prior to recruitment by WPPSI-III, [116] WISC-IV [117] or WAIS [118] (6) participants and parents willing to commit to the intensive therapy program for a 3-week period.

Exclusion criteria were (1) epilepsy (2) previous orthopaedic surgery or Botulinum Toxin A (BoNT-A) injection in the UL within 6 months prior to study entry.

The participants all volunteered and signed a written consent form once they had been correctly informed of the procedure.

Only six patients were enrolled for the pilot study because the aim of this study was to determine the validity and feasibility of the Tele-UPCAT project.

However, in order to detect a 1,40 effect size (value based on our preliminary data) at significant level of 0,05 and 80% power, 12 participants will be enrolled in the RCT study. This, according to the CONSORT guidelines, is the necessary number as 20% of possible drop-outs must be considered.

Full ethical approval was obtained from the Tuscany Paediatric Ethics Committee (169/2016, Protocol Version 5.0 of 10/11/2016) and any deviations from the protocol will be promptly notified to this Ethic Committee and applied only after its approval. The trial has been registered at http://www.clinicaltrials.gov (identifier NCT03094455).

Type of treatment

The Tele-UPCAT system was created thanks to through the close collaboration between the rehabilitation staff (child neurologists and child therapists) of IRCCS Fondazione Stella Maris and the biomedical engineers at the BioRobotics Institute of Scuola Superiore Sant'Anna. (Pontedera, Pisa).

The team of IRCCS Stella Maris designed 15 rehabilitation sessions to be carried out: one a day, for five working days, for three consecutive weeks. The daily session lasts sixty minutes for a total of 15 hours of treatment.

The participants undergoing the AOT intervention through the Tele-UPCAT system, watched three minutes of a video showing a child oriented goal-directed action and subsequently carried out the action proposed for three minutes. Each day three different actions were proposed twice.

All the exercises were shown from a first person perspective and were recorded using a GoPro against a black background so that the patient's attention would not focus on anything else other than the action itself. The videos showed only one hand and the forearm if the action was unimanual and both hands and forearms when the action was bimanual. Each video was edited to repeat the same goal-oriented action for three minutes.

A right-handed female actor performed the actions in the videos for participants with right UCP. For the left UCP participants the previous videos

were flipped 180° over if the same characteristic of the setting and hand movement were maintained, while the remaining videos were recorded anew.
A library of rehabilitation packages, composed of three different series of AOT exercises which were suitable to be carried out at home, was created by members of the rehabilitation staff viz. child neurologist and child therapists.
The exercises presented three levels of difficulty based on the HFCS grade of the patient's hand: one for HFCS 2-3, one for 4-5 and the last one for 6-8. A selection of commonly found toys and objects were used to create the sequences designed to cover unimanual and bimanual UL goal-directed actions. Each series, was composed of 15 sets of routine UL activities: 8 unimanual followed by 7 bimanual. Each set had a general common goal (e.g. drinking a glass of water). The participant was invited to perform three sequential tasks of increasing complexity (e.g. from picking up a bottle of water, to opening the cap, to pouring the water into a glass). As previously indicated, in order to grade the activities according to the range of capabilities, three series of sets were planned. The actions of each series had the same goal but the material and type of movement (i.e., range of movement, type of grasp) was customized in order to guarantee feasibility of the proposed activity while maintaining the same overall objective. For example, to grasp an object, children who had been classified with a House disability of 2-3 were given a Velcro bracelet to help them carry out the action. Other times the action requested in the series (HFCS 2-3, 4-5, 6-8) was slightly different, e.g. the participants were asked to tear apart a piece of Play-Doh using the affected hand or they were asked to use the affected hand to hold the Play-Doh and tear it apart using the non-affected hand.

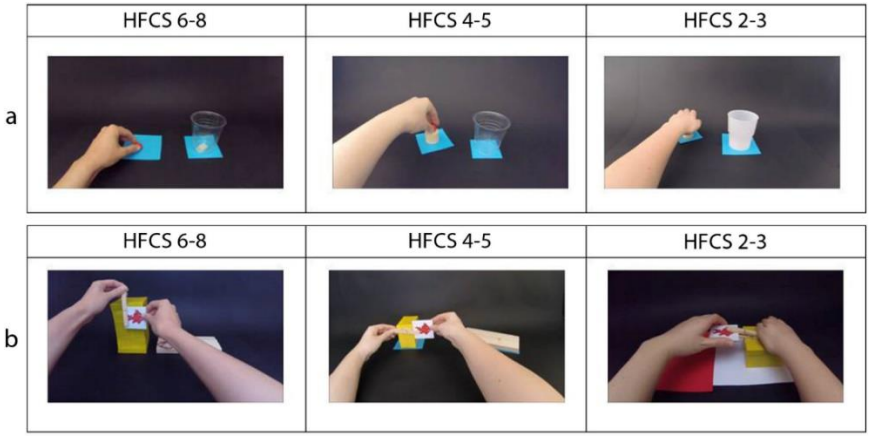


Fig.11 *Types of activities for different HFCS levels*

The patient observed the action for three minutes and subsequently repeated the action for three minutes. The observation and action phases were then repeated for a second time, each lasting 3 minutes as before.

A booklet explaining the exercises, the materials to use for each exercise, how to place material on the table etc. was compiled for the parents or tutor. Further useful information was also included.

The caregiver sat beside the patient and helped him/her out to fix the material on the table, focusing on the action it was shown and to perform the action as best as they could, and as similar as possible to the video.

The platform

Tele-UPCAT system (see Figure 4) was designed and built by the BioRobotics Institute of Scuola Superiore Sant'Anna (Pontedera, Italy).

It is a dedicated platform for delivering AOT and was designed to be user-friendly, by subjects at home in a playful setting with integrated smart features. All the participants had access to it.

The platform was designed and developed by integrating two different modules: the Observation Module (OM) for the presentation of AOT videos and recording of participant's attention and exercise execution, and the Motor Performance Module (MPM) for the execution of actions.

The OM used an all-in-one pc with 23" screen, a dedicated software and a video camera. The latter was a GoPro camera that was put on the pc screen and recorded a whole field, including the subject's face and hands and table with objects.

Moreover, a large switch that resembled a large red button was placed in front of the patient.

The Observation Module is important to determine whether the participant is looking at the monitor during the observation phase and has an overall view of the execution of actions.

A dedicated software, designed after an in-depth analysis of specific literature, was developed for guiding and motivating participants through the phases of AOT (observation followed by execution). In addition, the software was customized for the wide age range of participants providing an interactive game with an engaging story which was different for every day of training for school aged children, and a slide-show with a voice-over for adolescents and young people.

The various exercises for the children in the treatment were proposed as part of a mission that a strange and funny character had to perform to save the planets.

To start each session the patient had to press the red button when all the material was ready on the table.

The general architecture of the software is based on the following sequence: observation of a 3-minute video followed by execution of the same action for 3 minutes. Subsequently, the same video was replayed and then executed a second time. As stated before, a 60-minute session, including rest intervals, of three different goal-directed actions of increasing complexity are observed and imitated twice every day. At the end of each day the software terminated the session and automatically update it for the following day.

The Motor Performance Module (MPM) was mainly composed of a kit of exactly the same common objects and toys shown in the videos and two Actigraphs (wGT3X-BT, wActiSleep-BT) were worn, one on each wrist.

The Actigraphs were used to measure the asymmetry index during the treatment, during the assessment phase and in their daily lives.
All the measurements and videos could be observed remotely by the research group.

Experimental procedure

For each participant the most appropriate AOT rehabilitation package was chosen according to age and HFCS level. The engineers uploaded the package in the Observational Module (OM) while the therapists organised all the objects in a container and numbered them according to the training day they would be used. Moreover, a printed manual with instructions to set up and manage the system was provided along with guidelines for the different steps of the training. In addition, two Actigraphs (wGT3X-BT, wActiSleep-BT) were initialized for the recording period (3 weeks).
The ICT platform was delivered to the participant's home and a designated position for the table or desk of about 80x100 cm near a socket was chosen on which to place the ICT platform.
In the first few days of training a therapist and an engineer went to the patients' homes to define the location, install the system, explain the training and to teach the child and the caregiver how the treatment worked. The therapist also followed the first two training days to assist participants and their parents.
During the training sessions, each participant sat on a chair with both arms placed on a table in front of a platform positioned at about 1 m. A parent sat on the participant's affected side to prompt attention during task execution and assist if necessary.
The software guided the participant in the sequence of observations and executions.

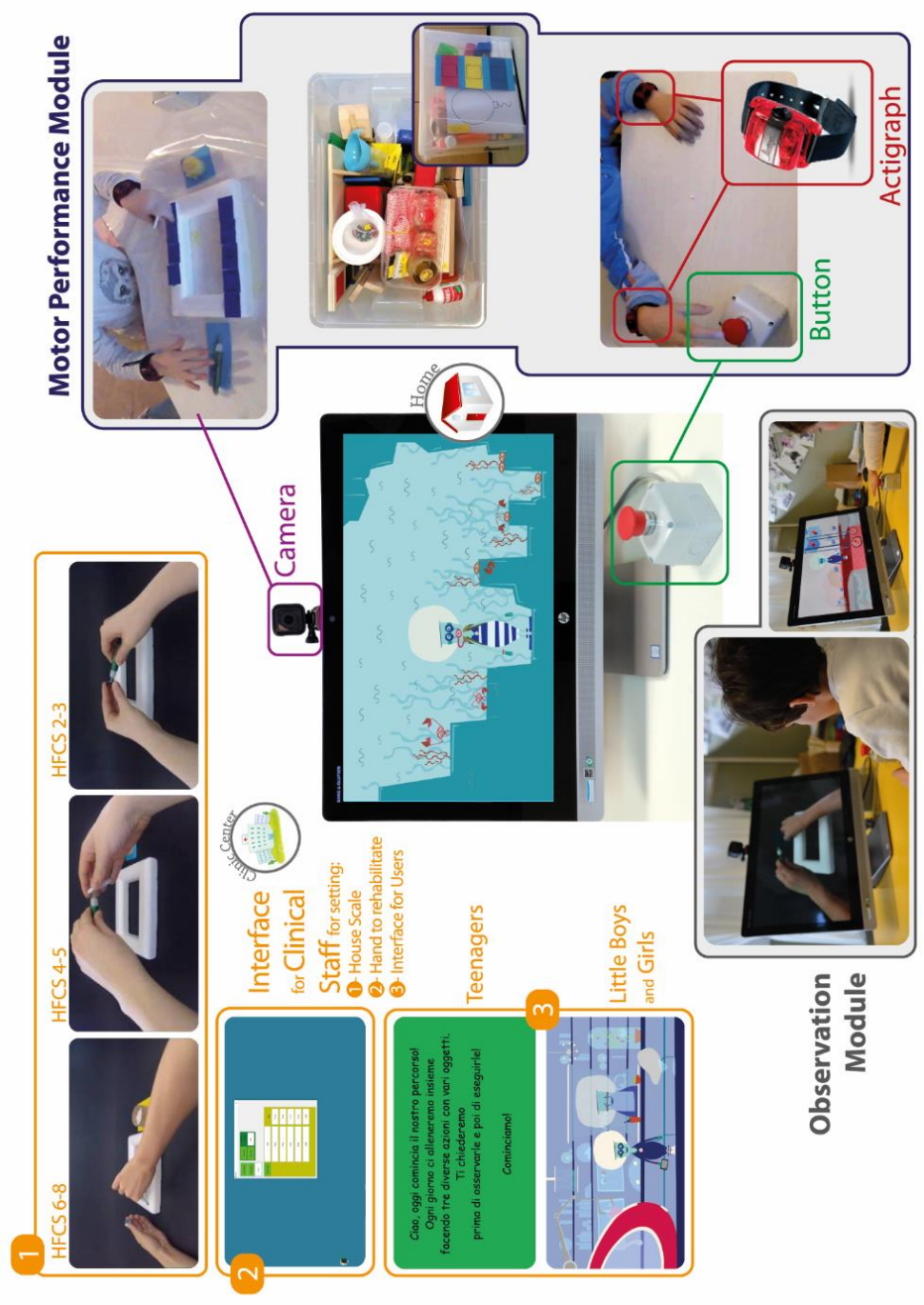


Fig.12 *Experimental procedure*

Outcome measures

Children participating in the study were classified according to HFCS, which assesses the function of the impaired hand in children with CP.

The primary outcome measure used in the study was AHA. The latest version of AHA 5.0 was used. AHA (Assisting Hand Assessment) is an assessment measure of the upper limb function during bimanual activities. It evaluates the spontaneous use of the assisting hand during semi-structured sessions. The sessions that are age appropriate, last for about 10-15 minutes and use specific toys or objects requiring bimanual handling.

Two forms of AHA were used according to the child's age: participants ranging in age from 6-12 years used the school kids form while the adolescent version (Ad-AHA) which uses a specific board game, was used with participants ranging from 13 to 18 years.

All AHA assessments were videotaped in a standardised manner and the subsequent scoring was carried out by a certified expert rater who was masked to group allocation and assessment order. The minimal detectable changes of AHA units reported in literature are 4,4 in children and 4,5 in adolescents. [119]

Of the many outcome measures that are currently being used for the assessment of the RCT, Box and Block Test (BBT) was chosen for this pilot study.

BBT is a quick (2-5 minute), simple and inexpensive test which measures unimanual dexterity in the activity domain. BBT is composed of a test box with a partition in the middle and 150 wooden blocks (25mm). The patient has to transport as many blocks as possible in 1 minute from one compartment to the other. Firstly, the patient is asked to perform the test with the unaffected hand and subsequently with the affected hand. The number of blocks transported by the affected hand in 1 minute is counted and considered for the main analyse. This test can be used for a wide range of populations from children to adults.[120]

The minimal detectable change in adults after a training of 2 weeks reported in literature is of 4 blocks. [44]

Both tests were scored blinded at T0 and after three weeks (T1).

Other secondary measures are included in the RCT trial: these are measures of unimanual capacity (Melbourne Assessment ((*MA-2),* and BBT) and bimanual daily activities at home and in the community (ABILHAND-Kids). Moreover, participation (PEM-CY) and quality of life (CP QOL) are also assessed. All assessment are performed at T0, T1, T1 plus, T2 and T3 unless otherwise indicated. Questionnaires are completed by parents and/or participants at home and if doubts occur, child neurologists or therapists are available to discuss face to face items which are not clear to them.

In this thesis only AHA and BBT scores were taken into consideration.

Engineers at Scuola Superiore Sant.Anna will analyse the data and calculate the asymmetry index.

Statistical analysis

The statistical analysis for the pilot study was performed using the means of the Statistical Package for Social Sciences (SPSS) 20.0. Due to the small sample size used, descriptive and non-parametric analysis (Wilcoxon test) was performed.

When the RCT trial ends the means and standard deviation of clinical outcome scores for both groups will be calculated to identify potential baseline differences between the groups. As a first step, normality of distributions will be verified by Shapiro-Wilk's test. In the RCT study a between-group difference for all the selected outcome measures will be evaluated at T0, by means of t-test for unrelated samples or non-parametric Mann–Whitney U independent sample test, for normal or non-normal distributed data, respectively. To test the first hypothesis, an analysis will be carried out to evaluate between-group differences for primary and secondary outcome measures at the primary endpoint (T1), compared with T0, by means of parametric or non-parametric tests for unrelated samples. The age, HFCS level, characteristics of usual care (in both groups) and supervision of caregiver will be analysed for further exploratory analyses. In addition, matched-pairs tests (t-tests or Wilcoxon) will be carried out in order to assess retention of effects

at follow-up periods (T1 or T1 plus, T2 and T3) relative to assessment before AOT training (T0 or T1 for experimental or waitlist group respectively). Bonferroni corrections will not be carried out in relation to the exploratory nature of the current RCT study. To detect if significant changes will correlate to HFCS levels, a correlation analysis between score changes after AOT training (T1 or T1 plus) and assessment before AOT training (T0 or T1) will be carried out. Finally, an exploratory within-group analysis will be performed for the waitlist group comparing changes during AOT with respect to those of the first standard care period.

5.3.2 Results

Six patients were selected with different HFCS grades. The patients mean age at T0 was 13,80 ± 3,60 years old with a range of 6,67-16,49 years. The patients all suffered from congenital hemiplegia.

According to the classification of Cioni et al. 1999, all the six children suffered from type 3 form of hemiplegia. Of the six patients three had a right affected hand and three had a left-affected hand. Four of the six patients were female and two were male.

The six children lived in six different towns: Lucca, Pistoia, Fauglia, Catania, Forlì and Carrara, so the home-based treatment was fundamental for them to be able to follow the intensive treatment.

All the six patients were assessed by a therapist at T0 and after three weeks of training.

Table 9. *Patients' data and scores*

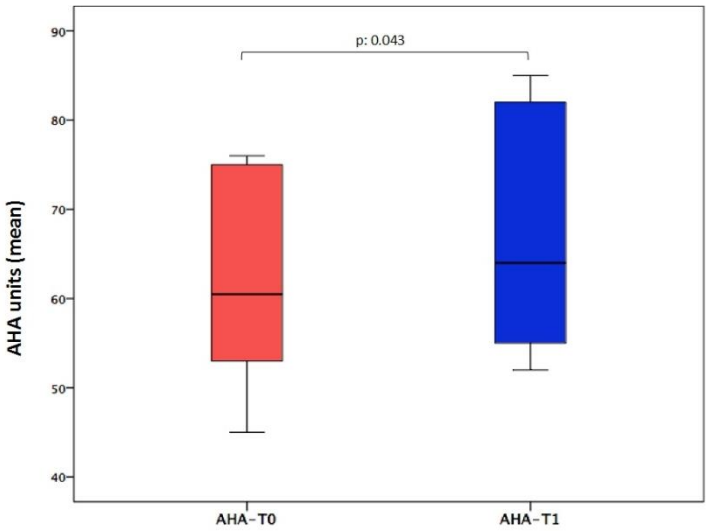
Subject	Age to T0 (Y)	Sex	Affected side	Type and timing [(a)]	HFCS	Outcome measures T0			Outcome measures T1			T1 - T0		
						AHA units	BBT affected	BBT non affected	AHA units	BBT affected	BBT non affected	AHA units	BBT affected	BBT non affected
1	9,64	M	R	Type 3	2-3	45	11	50*	52	13	50	7	2	0
2	9,30	F	L	Type 3	2-3	55	21	45*	55	29	59	0	8	14
3	6,47	F	L	Type 3	4-5	53	14	51*	57	22	61	4	8	10
4	12,48	F	R	Type 3	4-5	66	21	72	71	23	74	5	2	2
5	16,49	M	R	Type 3	6-8	75	35	80	85	38	85	10	3	5
6	13,80	F	L	Type 3	6-8	76	34	47*	82	36	57	6	2	10
Mean	13,80					61,7	22,67	57,5	67	26,83	64,33	5,33	4,17	6,83
Stand.dev.	3,60					12,6	9,97	14,71	14,38	9,41	12,80	3,33	2,99	5,38
z (T1-T0)									-2,02	-2,232	-2,032			
p (T1-T0)									0,043	0,026	0,042			

(a) according to the classification of Cioni et al;1999

* Out of the normal range according to Mathiowetz I; 1985

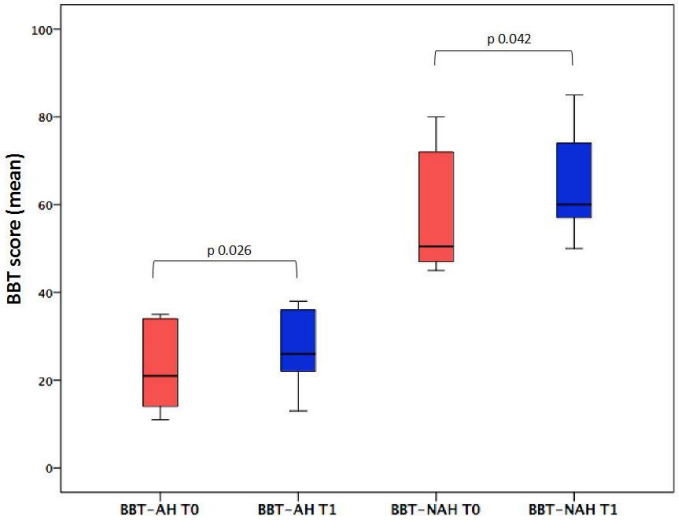
The patients showed an improvement of ability after training. The AHA difference T1-T0 mean was 5,33 ± 3,33 units with a range of 0-10. On the basis of the six patients, four had made a significant improvement in AHA units, whereas two had not: one totalize the same score at T0 and T1, and the other improved of 4 points.

Fig.13 *AHA at T0 and T1*



The mean of BBT scores in the affected hand was 4,17 ± 2,99 with a range of 2-8 from T0 to T1. Only two participants totalized a difference of 8 points, in the other patients the score did not change significantly. It was interesting to notice that the dexterity measured by BBT of the dominant hand also improved, the mean of the delta changes was 6,38 ± 5,38 with a range of 0 - 14. The improvement was significant in four patients. The four children had T0 scores that did not fall within the normal range for their age, according to those reported in literature by Mathiowetz (et al, 1985). Three of these four patients who prior to the training were out of range improved their scores at T1 significantly. After treatment two fell within the range of normality, and one fell closer to the normal range. Only one patient did not improve at all.

Fig 14 : *Box and Block Test affected hand (AH) and non-affected hand (NAH)*



When comparing the relationship between the results of AHA and BBT it can be noticed that two (#2, #3) of the six children improved at BBT but did not improve in AHA measures. The four children (#1, #4, #5 and #6) who improved more than the minimal detectable change at AHA did not improve at the BBT more than the minimal detectable change. However, two (#5, #6) of the children showed improvement in their non-affected hand.

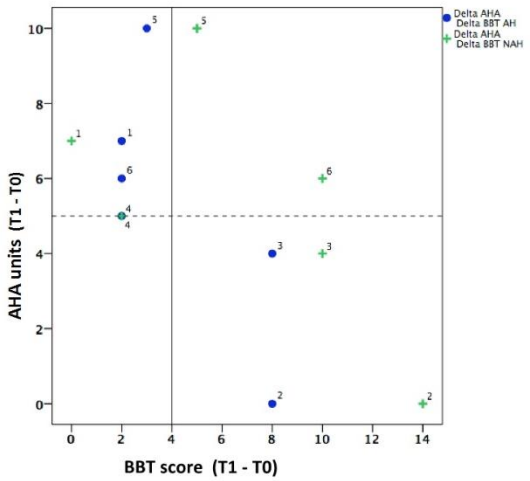


Fig.15 *Comparison of AHA units (T1-T0) results and BBT scores in the affected hand (AH) and in the non-affected hand(NAH)*

We decided to compare the means of the difference in the results obtained T1 and T0 dividing the children on the basis of the HFCS, because as previously reported, each group received different types of stimuli. The analysis reported in the following diagram shows that in the HFCS 2-3, the patients improved in the AHA units measure but not in BBT. The opposite happened in HFCS 6-8 where an increase in AHA units was more than the minimal detectable change but the increase in the affected hand in BBT was less than the minimal detectable change. This could be due to the bimanual ability requested in AHA which is not tested in BBT. However, more data is needed to test the validity of this descriptive data.

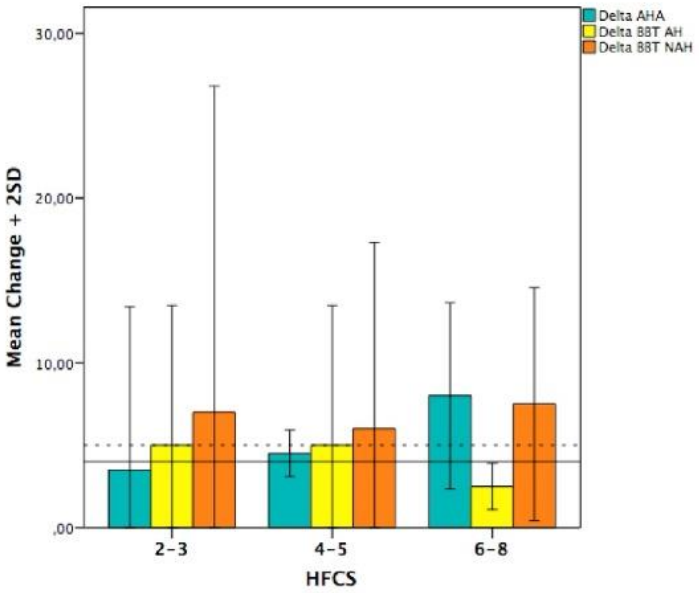


Fig. 16 *Comparison between AHA units (T1-T0) BBT scores (T1-T0) in the affected hand (AH) and in the non-affected hand (NAH) in the different HFCS groups.*

Finally, we compared the relationship between the scores of the BBT of the affected hand with those of the non-affected hand. An improvement was recorded in two patients in both scores i.e. of the affected and non-affected hand (#2 and #3). In two others (#5 and #6) the BBT scores improved in the non-affected hand only. Nonetheless in the remaining two (#1 and #4) the improvement was less than the minimal detectable change for both hands.

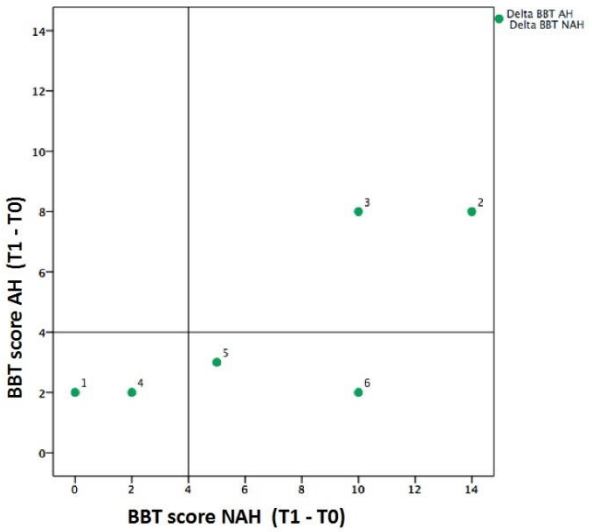


Fig. 17 *Comparison between the BBT scores (T1-T0) in the affected hand (AH) and in the non-affected hand (NAH)*

5.3.3 Discussion

The validity and feasibility of the study, which were the main aims of this pilot study, were verified. Moreover, the preliminary data of this pilot study showed an improvement of the hand ability domain of ICF after treatment with AOT. This improvement was reported both by AHA and BBT. However, the study was carried out only on six patients so this data have to be confirmed in the ongoing RCT trial.

We compared the results of this pilot study with those reported in literature.

The previous study where AHA was used as an outcome measure on children was by Sgandurra et al (2013). However, this study used logits to assess the patient at T0 and T1. As this study was carried out in IRCCS Stella Maris, we obtained permission from the authors to access the old database in order to calculate the delta units of the experimental group. This study was carried out in the rehabilitation hospital with a similar protocol. The difference between the units T1-T0 of this pilot study was 5,33 ± 3,33 whereas the delta units of the database of Sgandurra et al. were 4,67 ± 3,82. Due to this difference it is possible to hypothesise that a rehabilitative home-based treatment is better than treatment held in a hospital. It is reasonable to suppose that the child living in his/her own environment is more at ease and can concentrate better than in a rehabilitation centre. Moreover, the child and his/her family are under less stress as the child does not have to attend the rehabilitation centre on a daily basis.

Finally, a child in his own environment can experiment the effects of the treatment in his routine life and is therefore more motivated to follow the treatment seeing the progress made. However, to determine if there is a significant difference between a home-based therapy and a treatment carried out in a hospital further data will have to be analysed.

As regards BBT measures the pilot study was compared to the study by Kim (et al, 2014) which was carried out on children. Kim et al. found a mean

improvement after treatment of 7,88 ± 2,9 compared to 4,17 ± 3,00 of this home-based study. However, the studies differ in different aspects e.g in the sample group, of the 8 patients enrolled in Kim et al.'s study, 4 suffered from hemiplegia, 3 from diplegia and 1 from tetraplegia. Moreover, further data is necessary to compare the outcomes of the studies.

		Pilot study; 2017 [(a)]	Sgandurra et al.; 2013 [(b)]	Kim JY et al; 2014 [(b)]
AHA	delta	5,33 ± 3,33	4,67 ± 3,82	
	z	-2,023	-2,62	
	p	0,043	0,008	
BBT	delta	4,17 ± 2,99		7,88 ± 2,9
	z	-2,232		-3,188
	p	0,026		0,001

(a) based on Wilcoxon test
(b) basen on Mann–Whitney test

Table 10. *Comparison of the results of this pilot study with those reported in literature.*

As regards the improvement of BBT in the dominant hand the data suggests that children may have less dexterity compared to their peers. This type of AOT, thanks to the bimanual goal-directed actions, improves the dexterity of both hands. The dexterity of the dominant hand improved more if the range of T0 was less than normal for the child's age. However, the data available at this moment in time is not comprehensive so further studies are necessary to assess if there is a disparity between normal dexterity ranges and dominant hands in children with hemiplegia. Moreover, the study that calculated the normative data was carried out in 1985, and it may even be possible that coloured cubes, or timing of the study affected the results.

Future perspectives

The combined findings of this thesis and the analysis of previously published literature indicate that AOT can positively affect neuro-rehabilitation. Whilst AOT seems to improve the ability of children with congenital hemiplegia a wider RCT study is needed nonetheless. The ongoing RCT study will compare participants in the experimental group with standard care, and will also test the patients at T2 (after 8 weeks) and T3 (after 24 weeks). It will therefore be possible to verify if an improvement is found and maintained over a period of time.

Moreover, in the RCT study the actigraphs will test the asymmetry index during the participants daily lives, so it will be possible to see if this type of training helps the child to use the affected hand while carrying out daily routine activities.

More and more emphasis should be placed on home-based care and therapies for a number of reasons. Besides cutting costs this would not only increase efficiency and alleviate the workload of the hospital staff but would offer a wider population the opportunity to avail of treatment.

Children and young people in non-urban areas are usually at a disadvantage as often times they cannot access treatment easily due to the downsizing or closure of hospitals in their area.

If the ICT solution used in the Tele-UPCAT study was made available cost efficient rehabilitation programs could be developed. We suggest that the home setting might increase accessibility of rehabilitation to a larger number of children and young people with UCP (e.g. participants who live far from the clinical centres) with a large range of hand impairments (including also participants with HFCS level lower than 6) and older age (5-20 years instead of 5-15). Moreover, this type of approach would more than likely reduce family stress.

Further studies should be carried out to study the effects of AOT on children not only with hemiplegia, but also with other forms of Cerebral Palsy e.g. diplegia and tetraplegia. In fact, the UPCAT and Tele-UPCAT study which is

currently under way takes into consideration hemiplegic children only. In literature two different studies (Kim et al, 2014; Buccino et al, 2012) are on children with different types of CP but more specific sample sizes are needed for further analysis.

Finally, studies which have been carried out on adults with lower limb impairment are available in literature, but to date no studies on children have been published.

In the future we hope that home based AOT will become a tool to broaden the population who can gain benefit from rehabilitation treatment and that a larger number of patients will be able to take advantage of it. We also hope that further research will be conducted on different types of CP.

5.4 Usability and acceptability of the Tele-UPCAT system

Home-based therapies have been defined "*therapeutic activities that the child performs with parental assistance in the home environment with the goal of achieving desired health outcomes*." (Novak et al, 2007). They represent a way to facilitate the patient access to rehabilitation and enhance motivation. AOT can be brought into a home environment, especially thanks to technology, leading to a tele-rehabilitation practice. Through telecommunication devices (telephone, videophone, emails) and monitoring, patients can carry out, together with parents and/or caregivers, an intensive rehabilitation therapy directly into their home. Given the economic and energy advantage that such a therapy brings, other issues must be investigated: the usability and acceptability of a technology-based therapy into the home environment. The requirements and expectations of the user must be considered to ensure the success the system. The use of technology in the rehabilitation domain changes the interaction between patients and rehabilitation and can limit acceptability of new technology. Usability is defined as "the extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use", while acceptability as "demonstrable availability to use technology (Davis, 1993), and the way people perceive, accept, and adopt technology use" (Fred et al, 1989).

5.4.1 Methods

The aim of this pilot study was to assess the usability and acceptability of a new technology for home-based rehabilitation, named Tele-UPCAT (Upper Limb Children Action Observation Training). The rationale of this study is represented by the need to analyse different factors that affect the introduction of technology for home-based rehabilitation and the perceptions of the end-user on a treatment provided directly into the home environment and on the use of ICT for the therapy on a daily basis. The patient's needs are central in the development of health-care technology, especially those aimed to a home-

environment use. The results of this pilot are crucial to understand the issues and barriers, and the strength points as well, of a home-based therapy involving technology, to better meet the patient's expectations and desires for future home-based treatments.

Study design

The Tele-UPCAT RCT trial, which is currently underway, is an exploratory randomized, allocation concealed (waitlist-controlled), and evaluator blinded clinical trial. The study compares the AOT home-based intensive rehabilitation programme to standard care in children and adolescents with UCP. (Methods are showed above in the study protocol).

This pilot study was carried out in the Tele-UPCAT project RCT.

Participants

15 patients (children and adolescents) who have already performed the AOT training were selected to fill in the questionnaire about the usability and acceptability of the treatment. The age range was between 7,50 and 18,75 years, the mean age was 11,95 ± 3,46 (M:F=7:8, 10 right UCP, 5 left UCP). Sample characteristics are showed in the table 1.

N.	AGE	SEX	HFCS	AFFECTED SIDE
1	8,90	F	4-5	L
2	12,00	M	4-5	R
3	13,90	F	6-8	L
4	18,75	M	6-8	R
5	16,50	M	6-8	R
6	15,16	M	6-8	R
7	8,08	M	4-5	L
8	7,50	M	6-8	L
9	12,50	F	4-5	R
10	9,75	F	4-5	R
11	15,33	F	6-8	R
12	9,16	M	2-3	R
13	13,50	F	6-8	R
14	8,90	F	4-5	R
15	9,33	F	2-3	L

Table 1. sample characteristic

The questionnaire (Fig. 1)

To assess the usability and acceptability of the whole therapy, a 32-query questionnaire was composed. It is specific for the Tele-UPCAT program, but it is based on the standard definitions of usability and acceptability and the factors generally evaluated in this kind of assessment (fig. 2).

Semi-Structured Interview on the Usability and Acceptability of the Tele-UPCAT system

Customisation of the exercises						
1	Were the exercises suitable for you (materials, type of actions, size...)?	no, no one (1)	2	3	4	yes, all of them(5)
2	Were the exercises difficult for you (required perfomances)?	no, no one (1)	2	3	4	yes, all of them(5)
3	Did you notice an increasing difficulty (from the easiest the first day to the most difficult)?	no, no one (1)	2	3	4	yes, all of them(5)
4	Were the exercises similar to other typical activities/daily life actions?	no, no one (1)	2	3	4	yes, all of them(5)
5	Do you think the AOT has improved your abilities?	no, not all(1)	2	3	4	yes, a lot (5)
6	Do you think watching the videos twice was helpful?	no (1)	2	3	4	yes (5)
7	Did you like the unimanual exercises?	no, not all (1)	2	3	4	yes, a lot (5)
8	Did you like the bimanual exercises?	no, not all (1)	2	3	4	yes, a lot (5)
Acceptability of the Tele-UPCAT system at home in daily life						
9	Did you like to do the training at home?	no, not all (1)	2	3	4	yes, a lot (5)
10	Did you like to perform exercises without a therapist?	no, not all (1)	2	3	4	yes, a lot (5)
11	If not, did you reorganise your space at home?	yes, a lot (1)	2	3	4	no, not at all(5)
12	Do you think the whole system was bulky?	yes	2	3	4	no
13	Was the commitment (1 hour a day) feasible?	no, not at all (1)	2	3	4	yes, very much (5)
14	Did you need to change your normal routine?	yes, very much (1)	2	3	4	no, not at all(5)
15	Did you like to have a daily commitment?	no, not at all (1)	2	3	4	yes, very much (5)
16	Do you think you could have continued?	no, not at all (1)	2	3	4	yes, very much (5)
Required Effort						
17	Do you think the exercises had some negative aspects? (characteristics, duration, difficulty...)	yes, all of them (1)	2	3	4	no, no one (5)
18	Did you feel uncomfortable or painful during/after the exercises?	yes, always (1)	2	3	4	no, never (5)
19	Did you enjoy wearing the Actigraphs?	no, not at all (1)	2	3	4	yes, a lot (5)
20	Were you uncomfortable while wearing the Actigraphs?	yes, a lot (1)	2	3	4	no, not at all(5)
21	Did you wear the Actigraphs for the majority of the time of the training (night and day)	no, never (1)	2	3	4	yes, always (5)
22	Did you remember how to wear them every time?	no, never (1)	2	3	4	yes, always (5)
23	Did you remember to fill in the diary?	no, never (1)	2	3	4	yes, always (5)
24	Was the management of the whole system burdensome?	yes, a lot (1)	2	3	4	no, not at all(5)
Suitability of the Manual/Software						
25	Was the manual clear enough?	no, not all (1)	2	3	4	yes, a lot (5)
26	Did you have any difficulties in finding/preparing the material?	yes, always (1)	2	3	4	no, not at all(5)
27	Were the instructions for the managing of the system (go pro, objects…)complete and clear?	no, not at all (1)	2	3	4	yes, a lot (5)
28	Did you like Ubi/Slides?	no, not at all (1)	2	3	4	yes, a lot (5)
29	Do you think there is something that should be changed?	yes, a lot (1)	2	3	4	no, no one (5)
30	Was using the computer difficult?	yes, a lot (1)	2	3	4	no, not at all(5)
31	Did you have technical issues?	yes, a lot (1)	2	3	4	no, not at all(5)
32	Did you ever need technical assistance?	yes, a lot (1)	2	3	4	no, not at all(5)

Fig. 1 the questionnaire

The questions are divided in 4 sections composed of 8 queries, that analyse crucial items of the therapy, from the characteristics of the exercises to the acceptability of the Actigraphs and enjoyment of the whole therapy.

The groups are:

- "Customisation of the exercises": how the patients perceived the exercises as personalised following their abilities and needs, how difficult they were perceived, their preference about the kind of exercise.
- "Acceptability of the Tele-UPCAT system at home in daily life": how patient coped with a technology system directly installed in their home and the commitment to the therapy in everyday life.

- "Required effort": whether the patient perceived the whole treatment as tiring and strenuous and the use of Actigraphs every day as bothersome.
- "Suitability of the manual/software": whether the manual was complete and clear in delivering the instruction to use the system and if the software (including the program) was enjoying and easy to use, without technical issues.

These four areas and the questions have been planned relative to the attributes of usability and acceptability, as shown in figure 2.

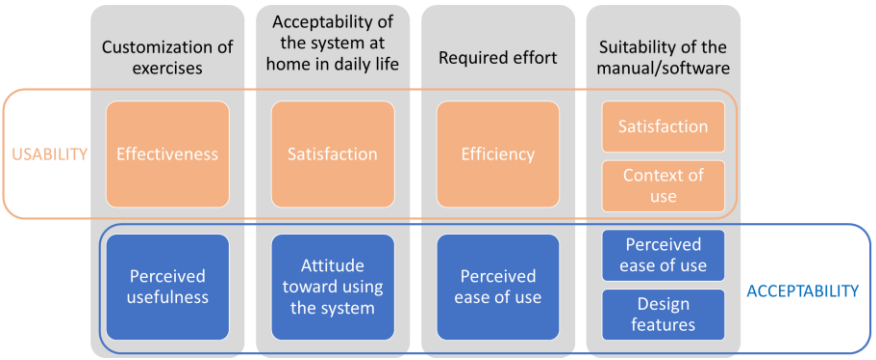


Figure 2 Links Between the Questionnaire groups and the attribute of Usability (ISO9241-11) and Acceptability (TAM)

The questionnaire was delivered soon after the end of the AOT therapy period (T1 or T1 plus) in a face-to-face interview with the patient, with or without the help of the parents or sent via email if the participant had already finished the 3-weeks program, through a version specifically composed and uploaded on Google Modules, with the same questions.

The face-to-face interview allowed the interviewer to explain the queries when necessary and the participant to give their perception more freely. The Google questionnaire, indeed, was appreciated because easy and quick to fill in, with the possibility to add personal thoughts as well. It was preferred also for adolescent and young adult participants. Young children were helped with a smileyometer to facilitate them in expressing their feelings (fig. 3).

Fig. 3 smileyometer

Participants were asked to supply personal thought on every item of the questionnaire, to have a wider view on the patient's insight.

Data collection

The questionnaire was delivered immediately after the end of the AOT therapy period (T1 or T1 plus) in a face-to-face interview with the patient, with or without the help of the parents. Alternatively, if the participant had already finished the 3-weeks program, the same questionnaire was sent via e-mail, through a version specifically composed and uploaded on Google Modules. The face-to-face interview allowed the interviewer to explain the queries when necessary and the participant to give their perception more freely. The Google questionnaire, indeed, was appreciated because easy and quick to fill in, with the possibility to add personal thoughts as well. It was preferred also for adolescent and young adult participants.

5.4.2 Results

All the 15 children and adolescents gladly accepted the interview or the on-line questionnaire, adding also some personal comments.

Every section of the questionnaire can range from a minimum of 8 (strongly negative) to a maximum of 40 (strongly positive), for a total of 160 for the whole questionnaire, indicating the greater level of acceptability and usability. In general, all participants showed a high level of acceptability and usability, with total scores all above 116 (72,50%). Regarding the 4 sections, "customisation of the exercises" is the one with lowest scores with a mean score of 31,93±4,59 (range 22-38); then acceptability of the Tele-UPCAT system at home in daily life has mean scores of 32,33±3,22 (range 27-38); required effort presents mean scores as 33,26±3,15 (range 29-40) and suitability of the manual/software 37,53±1,64 (range 34-40).

There were no differences in the total scores of answers related to the sex of subjects, the ages or the HFCS level (not considerable because of the small sample for the HFCS level 2-3, only 2 children, and the heterogeneity of the quantity of subjects in the three groups) or the different version of the software age-related (Ubi or slides).

For these reasons, we decided to present the data dividing the whole sample in two groups based on the age and the used software for the training: Ubi and its missions in the galaxy for children aged less than 12 years and a power point presentation with simple slides and voice guide for adolescents (>12 years). In our opinion, this could have highlighted some different opinions between these two groups, as seen from the spontaneous comments collected. Results are presented below.

The general opinion of the interviewed sample was globally positive and there were not significant differences within groups. This is the first interesting data, because it means that the effort of creating a customised training, which takes care of sex and age preferences, was recompensed: all the opinions were coherent and showed appreciation and positive feedbacks.

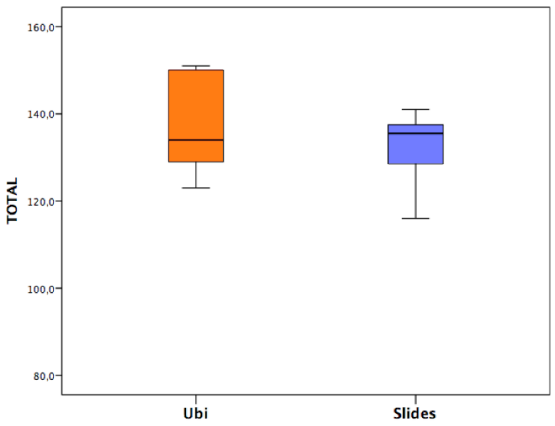


Fig. 4 total answers of the 2 groups

When observing the total score of answers, the median values were similar in the two groups, but the scores distribution showed higher values in younger subjects.

Customisation of the exercises

This section showed the lowest scores. The median values between the two groups are similar but the distribution of the answers is towards higher values in younger subjects and lower in older.

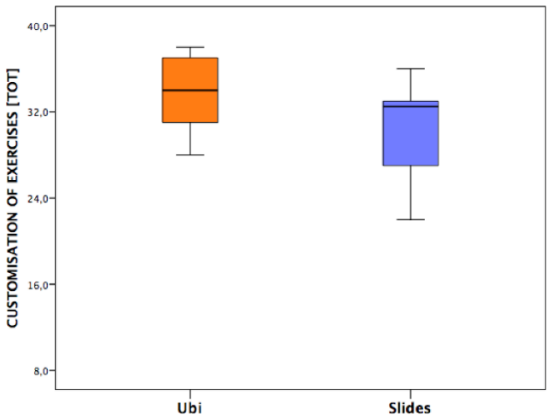


Fig. 5 answers of section "Customisation of the exercises"

Acceptability of the Tele-UPCAT system at home in daily life

In this section the younger children showed more variability in their answers than adolescents, who gave more similar answers.

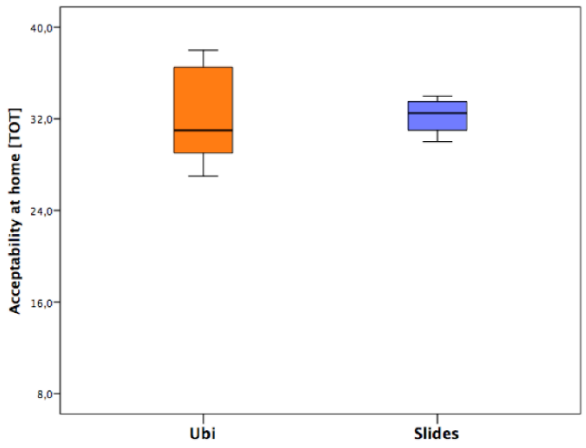


Fig. 6 answers of section "Acceptability of the Tele-UPCAT system at home in daily life"

Required effort

This section presented similar median values between the two groups, but the distribution is different: in fact, for children aged less than 12 years the required effort is perceived as feasible and the tendency of the scores is towards higher values, while for adolescents the distribution is equal around the median value and globally lower than children's scores.

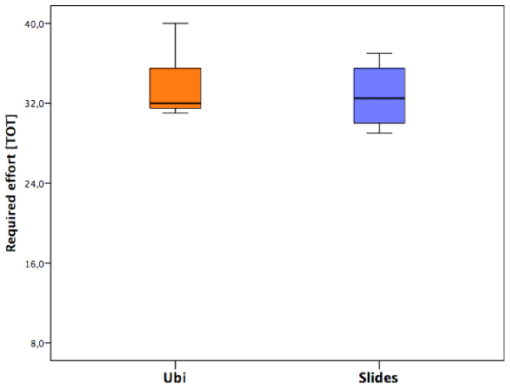


Fig. 7 answers of sections "required effort"

Suitability of the Manual/Software

In the last section, one positive aspect is that almost no one needed technical assistance or encountered technical issues during the training and this supports the stability of the system and its consequent appropriateness for home trainings.

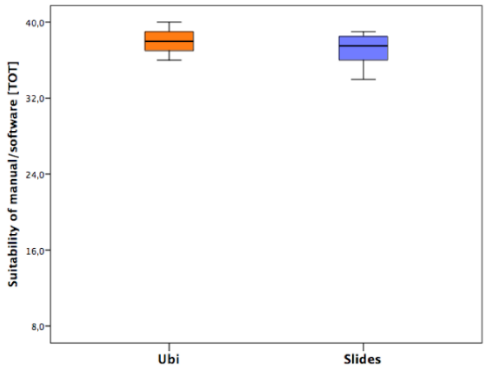


Fig. 8 answers of sections "Suitability of the Manual/Software"

5.4.3 Discussion

The general opinion of the interviewed sample was globally positive and there were not significant differences between the groups. This is the first interesting data, because it means that the effort of creating a customised training, which takes care of sex and age preferences, was recompensed: all the opinions were coherent and showed appreciation and positive feedbacks.

When observing the total score of answers, the median values were similar in the two groups, but the scores distribution showed higher values in younger subjects. This could be explained because children lived the training in a playful way and performed the exercises with a more motivational software (Ubi had to achieve missions). This result could also be due to the presence of their parents during the treatment. On the other hand, some adolescents demonstrated to like less the and a deeper analysis of the answers of the 4 sessions gave a clearer view on this aspect.

Customization of the exercises

This section showed the lowest scores and the median values between the two groups are similar, but the distribution of the answers is towards higher values in younger subjects and lower in older.

Perhaps, this could be explained by the fact that the objects selected for the exercises were more suitable for younger children in terms of sizes and features. In fact, the pool of objects and toys has been originally selected to be suitable for a wide range of ages, thinking about making them more compliant for younger children. However, this could mean that sometimes the toys resulted not suitable or excessively small for adolescents; moreover, children aged less than 12 years probably preferred to play with toys while adolescents could have appreciated something more appropriate for their age and interests. This could clarify the different opinions between the two groups.

In addition, the exercises maybe resulted poorly engaging and boring for adolescents because of their repetitiveness. Despite this, some interesting comments came from adolescents' opinions. Although they have less appreciated the training, they were more conscious about its relevance, as reported by an example of their personal comments as: " I have used some movement of the training in daily life", "I have used the hand to do certain movements I didn't know I was able to do", "Now I think about the right movement I've seen during the training and I can do it better", "I'm more aware of my hand now".

Acceptability of the Tele-UPCAT system at home in daily life

In this section the younger children showed more variability in their answers than adolescents, who gave more similar answers. Focusing on the specific answers, we noticed that an intensive training at home, which means dedicated time and space for subjects and their family, is feasible but requires organization of home spaces and daily activities. In the children group, answers' variability is probably related to the week routine: there were subjects who had many therapies and other family activities during the days and consequently attending the daily training was more demanding; on the other hand, some others were less busy, and the steady commitment of the training represented an appreciated routine.

Required effort

This section presented similar median values between the two groups, but the distribution is different: in fact, for children aged less than 12 years the required effort is perceived as feasible and the tendency of the scores is towards higher values, while for adolescents the distribution is equal around the median value and globally lower than children's scores.

The questions of this section were more focused on the use of Actigraphs and the level of difficulty of exercises. From the free comments in the questionnaire, it emerged that the wristbands were not quite comfortable, and

some adolescents reported itch or bother and, in some cases, also embarrassment while wearing Actigraphs in social contexts (school, parties, free-time).

Furthermore, the difficulty in performing some actions emerged more in adolescents. It has been explained because they were more aware of exercises movements features and they reported pain or complaint due to the frequency of the requested movements (done with two intervals of 3 minutes each).

Several subjects found the duration of the videos excessive, thus the evaluation of the exercises as boring. They understood the meaning of the two observational intervals and they demonstrated good levels of attention to the videos, but all of them judged the situation as boring and the videos a little bit too long.

Suitability of the Manual/Software

In the last section, one positive aspect is that almost no one needed technical assistance or encountered technical issues during the training and this supports the stability of the system and its consequent appropriateness for home trainings.

The manual and software were overall considered as clear and complete. Especially the children found the game of their specific software amusing and fun. The slight difference in the answers is basically related to the already reported different software features of the two groups.

5.5 Final remarks

Action Observation Therapy is a new innovative tool which, according to literature, brings about a significant improvement in activity and body function in the child and adults. As the most common type of treatment suggested is the observation of videos followed by the reproduction of the goal-directed action, a further standardization of the type of perspective used to record the videos and the length and intensity of the treatment should, however, be carried out. In the majority of studies, the intervention is focused on the upper limb. It would be desirable if in the future a larger number of studies investigated the lower limb function in children and identified a homogeneous population, e.g. the effect of AOT in a sample of patients suffering from diplegia or tetraplegia, rather than investigate a larger sample including children suffering from hemiplegia. Since the first AOT treatment was carried out by Ertelt in 2007, the number of studies on adults and children has increased. It is for this reason that a new type of home-based AOT has recently been proposed, thus broadening the population to whom this type of rehabilitation can be offered.

The combined findings of this chapter and the analysis of previously published literature indicate that AOT can positively affect neuro-rehabilitation. Whilst AOT seems to improve the ability of children with congenital hemiplegia and a wider RCT study is needed also. The ongoing RCT study will compare participants in the experimental group with standard care and will also test the patients at T2 (after 8 weeks) and T3 (after 24 weeks). It will therefore be possible to verify if an improvement is found and maintained over a period of time.

Moreover, in the RCT study the actigraphs will test the asymmetry index during the participants daily lives, so it will be possible to see if this type of training helps the child to use the affected hand while carrying out daily routine activities.

More and more emphasis should be placed on home-based care and therapies for a number of reasons. Besides cutting costs this would not only increase efficiency and alleviate the workload of the hospital staff but would offer a wider population the opportunity to avail of treatment.

Children and young people in non-urban areas are usually at a disadvantage as often times they cannot access treatment easily due to the downsizing or closure of hospitals in their area.

If the ICT solution used in the Tele-UPCAT study was made available cost efficient rehabilitation programs could be developed. The home setting might increase accessibility of rehabilitation to a larger number of children and young people with UCP (e.g. participants who live far from the clinical centres) with a large range of hand impairments (including also participants with HFCS level lower than 6) and older age (5-20 years instead of 5-15). Moreover, this type of approach would more than likely reduce family stress.

In addition, other interesting preliminary data were those about patients' and families' opinions about the home training. Data from the pilot exploratory study based on the questionnaire of the feasibility of the training, showed that patients and family demonstrated to be satisfied about the home training and confirmed the usability of the Tele-UPCAT system and Actigraphs and the acceptability of the training at home.

This is an important starting point on which work on for future studies and suggests that home training could be a valid option both from clinicians and families point of view.

Further studies should be carried out to study the effects of AOT on children not only with hemiplegia, but also with other forms of Cerebral Palsy e.g. diplegia and tetraplegia. In fact, the UPCAT and Tele-UPCAT study which is currently under way takes into consideration hemiplegic children only. In literature a few studies are on children with different types of CP but more specific sample sizes are needed for further analysis.

Finally, studies which have been carried out on adults with lower limb impairment are available in literature, but to date no studies on children have been published.

In the future, perhaps home based AOT will become a tool to broaden the population who can gain benefit from rehabilitation treatment and that a larger number of patients will be able to take advantage of it.

Data reported in this Chapter are based on the following papers:

- Sgandurra G, Cecchi F, **Beani E**, Mannari I, Maselli M, Falotico FP, Inguaggiato E, Perazza S, SicolaE, Feys H, Klingels K, Ferrari A, Dario P, Boyd R, Cioni G *Tele-UPCAT: Study Protocol of a Randomized Controlled Trial of a Home-Based Tele-monitored UPper Limb Children Action Observation Training for Participants with Unilateral Cerebral Palsy*, submitted
- Sgandurra G, Buchignani B, Iacono O, Inguaggiato E, **Beani E**, Sicola E, Feys H, Cioni G, *Action observation for upper and lower limb rehabilitation in hemiplegia: a systematic review*, submitted
- **Beani E,** Buchignani B, Cecchi F, Mannari I, Sgandurra G, Cioni G *The tele-upcat project RCT: a pilot study on a sample group of children with unilateral cerebral palsy*, submitted
- **Beani E**, Avola M, Sgherri G, Menici V, Sgandurra G, Cioni G *Acceptability and usability of tele-UPCAT system for Action observation training in children with Unilateral Cerebral Palsy*, submitted

CHAPTER 6

SUMMARY AND CONCLUSIONS

Neurodevelopmental disorders encompass a group of clinical heterogeneous conditions with onset in the developmental period and their early and accurate detection and intervention are crucial for aim to the best outcome.

The research is constantly search for new solutions for the standardization and quantification of the assessment procedures and rehabilitative treatments and this is process is giving birth to new tools.

This is possible mainly thanks to the joined work between engineering and medicine, which represents the field in which these tools are developed.

The work aimed to design, build, test and validate new technological tools is hard and requires efforts, but it has several advantages.

First, the quantitative assessment provides quantitative data which allow to deeply investigate the target function or behaviour and complete the traditional assessment. This aspect increases the reliability of data and overcomes the necessity of a higher trained assessor, standardizing and homogenizing the assessments.

This process needs an accurate optimization of the acquisition of data and a consequent analysis and interpretation of results and could not be easy from the beginning or highlight the unexpected limits, as seen for the assessment of evaluation of gaze capabilities of infants. CareToy C and the eye tracker for the visual assessment of young infants showed its utility and also its complexity of use, not in absolute but in relation to the challenging purpose. This quantitative assessment is a strong stimulus and represents an innovative starting point for further improvements as the on-line control, which could represent the next aim.

CareToy H, despite its variety of integrated modules which could complicate its applicability and use, demonstrated its validity and suitability to purposes, both for assessment of and intervention.

Moreover, CareToy H open the new frontier of the rehabilitation at home: thanks to the tele-rehabilitation, infants perform their program with their families at home, while the clinical staff follows their training and customize their next activities. This is a real turning point of the early intervention for

infants with neurodevelopmental disorders and a new model of enriched environment and presents many advantages as i) the highly customization of the training, ii) the central role of the family, highly involved in the training iii) the possibility to increase the intensity of the training, iv) the increasing knowledge of the aspects of the early development of infants. Of course, the CareToy will be upgraded and some modifications could additionally improve it and in the meanwhile the direction should be to test its feasibility also for other type of neurodevelopmental disorders in addition to infants at risk for Cerebral Palsy and infants with Down syndrome, already tested. Starting from this point, in a new project of IRCCS Stella Maris, CareToy is used together with a modular postural system for expand the use of CareToy also in a sample of infants with worse clinical situations, as brain lesions, which need higher postural supports and stabilization. This study will confirm the modularity and versatility of the CareToy H.

Finally, focusing on the assessment and rehabilitation of older children with Unilateral Cerebral Palsy, the AOT is one of the newest approaches and its combination with ICT will demonstrate again and more accurately its effects. In fact, technological devices as inertial sensors are demonstrating their usefulness and reliability in measuring upper limbs use and asymmetries and in literature there are some studies in which they are used for the assessment of upper limb movements; after their validation, they will represent a more quantitative and accurate evaluation of the use of the dominant and non-dominant hand and arm in relation to each other and the organization of bimanual activities.

For this reason, inertial sensors will be used in AOT programs for quantitatively measure effects of training and a further perspective could also be their use for analyse trends related to the development or progresses of an intervention, so not to limit their use for a single assessment but also for a long monitoring of upper limbs use.

Stating the increasing need of quantitative tools for assessment and rehabilitative purposes, the union between medicine and engineering could

overcome this lack. We are just entered in the period of change in this sense and there's a long road ahead, but this work of thesis aims to represent a milestone for the quantitative measures of spontaneous and intervention-induced perceptual-motor development in children with neurodevelopmental disorders.

CHAPTER 7

ACKNOWLEDGEMENTS

Desidero prendermi tutto lo spazio necessario per scrivere i ringraziamenti di questa tesi e voglio farlo in italiano per provare ad esprimere al meglio ciò che voglio dire…anche se non è detto che trovi comunque le parole giuste.

Grazie al Professor Giovanni Cioni, il mio Tutor, per aver creduto in me come "ricercatrice" prima ancora che io potessi capire cosa significava e per avermi incoraggiata e valorizzata a partire dalla Laurea Triennale fino ad oggi, dandomi la possibilità di lavorare in una struttura così stimolante quale è la Stella Maris.

Un grazie enorme va a Giusy, la Dr.ssa Sgandurra. Non ho difficoltà a dire che mi hai insegnato praticamente tutto quello che so sulla ricerca, mi hai fatta crescere davvero tanto, responsabilizzandomi sempre di più. Grazie per avermi considerata sempre per ciò che so fare, per essere stata sempre sincera e per avermi sempre e dico sempre voluta accanto a te. Grazie per tutti i "brava" che mi hai detto, ma grazie soprattutto per quelli che non mi hai detto. Grazie per avermi fatto sentire il tuo progetto una cosa anche un po' mia, so quanto è importante per te e ne sono davvero fiera e orgogliosa. È bello essere il tuo braccio destro!

Grazie Elisa Sicola e Silvia Perazza, per avermi avuta sempre tra i piedi durante le valutazioni, per avermi presentata a tutti i vostri bambini con il sorriso, per avermi dato fiducia, per aver sempre trovato un minuto per me.. e per aver riso insieme su tutte le cose che vi ho raccontato!

Elisa, grazie per avermi trasmesso la passione per le cose che facciamo e per essere, spesso inconsapevolmente, un vero modello da seguire. Ti stimo davvero moltissimo.

Grazie Valentina Menici, una collega e anche un'amica che non si è mai tirata indietro dal darmi una mano e che ha condiviso con me i momenti di caos e follia, grazie perché quando serve sai chiedere scusa e dire grazie. Grazie per non essere impazzita ed essere diventata pignola e perfezionista!

Grazie Bianca, può sembrare che tu abbia avuto un ruolo marginale, ma non è così. È stato bello lavorare insieme e ritrovarsi, grazie per la tua pazienza e grande professionalità, sono certa che sarai un'ottima Neuropsichiatra

infantile. Grazie a chi mi è stato vicino non solo durante questo percorso, ma anche prima (e spero poi!), e che in questi ultimi 3 anni mi ha sopportata ogni volta che ho detto che non ce l'avrei fatta: Simona, Ilaria, Jessica, Silena, Sara, Tanja, siete delle persone meravigliose, che mi fanno bene al cuore. Grazie per la vostra amicizia!

Grazie a chi nella mia vita c'è a prescindere, alle amiche di sempre, a chi c'è a periodi, a chi non sa più di esserci, a chi vorrebbe esserci…ognuno di voi fa parte in qualche modo di quello che sono e oggi voglio ringraziarvi.

Potrei dilungarmi ancora moltissimo, perché se fino a ora ho sempre vissuto gli eventi importanti come inizi, stavolta sento questo traguardo come una sorta di arrivo e sento che avrei voglia di scrivere ringraziamenti per tutte le persone che ho attorno. Sarà che concludo un percorso professionale molto importante e credo che (almeno per un po'!) non avrò il tesserino da studentessa, sarà che sento la responsabilità di avere quasi 30 anni, sarà che ho imparato a viaggiare anche da sola, sarà forse che dopo tanti anni a pensare a cosa farò da grande adesso un po' "grande" mi ci sento…

Concludo con i ringraziamenti più importanti, più belli, più sinceri, che vorrei tanto fare tutti i giorni.

Grazie alla mia famiglia, la mia forza e il mio porto sicuro. Non è scontato volersi bene e si sa che la famiglia ci capita, non la scegliamo. Ecco, secondo me probabilmente noi ci saremmo scelti comunque!

Grazie mamma e grazie babbo, siete sempre con me in ogni cosa che faccio, siete due genitori unici. Sentirmi amata e sostenuta mi ha permesso di crescere libera e felice, mi ha fatta essere determinata e mi ha fatta sentire sicura di poter fare tutto, perché mi avete insegnato che posso sempre contare su di me come su voi e so con certezza che qualsiasi cosa posso o voglio fare, voi ci siete.

Grazie babbo, per essere ancora oggi un riferimento fondamentale e un modello da seguire. I valori che mi hai trasmesso sono sempre chiari dentro di me, nella vita e anche nel mio lavoro e non hai idea di come mi sento quando a volte questo emerge e mi dicono che ti somiglio.

Grazie mamma, per la mamma che sei, grazie per tutti gli “sbagli” che fai, frutto di tutto l’amore che provi per me e Michele, perchè dimostrano quanto ti impegni in ogni modo per darci sempre il massimo. Grazie per aver imparato a lasciarmi fare anche un po’ come mi pare.

Grazie Michele, il nostro legame è indescrivibile e diventa giorno dopo giorno più forte. Sei parte inscindibile di me, mi fai stare con i piedi per terra senza mai farmi smettere di sognare e sei sempre assolutamente e totalmente sincero con me. Grazie per esserci in mille modi diversi, che mi piacciono tutti.

Grazie Marino, per regalarmi ogni giorno la tua parte più bella, la tua essenza, il tuo amore. Grazie perché credi tanto in me. Grazie perché sai sempre come prendermi, hai la capacità di dire o fare sempre la cosa giusta. Grazie per aver creato insieme il nostro noi ed esserci.

CHAPTER 8

REFERENCES

- Andersen GL, Romundstad P, De La Cruz J et al. *Cerebral palsy among children born moderately preterm or at moderately low birthweight between 1980 and 1998: a European register-based study*. Dev Med Child Neurol 2011;10:913-9. doi: 10.1111/j.1469-8749.2011.04079.x
- Arnould, C., Penta, M., Renders, A., & Thonnard, J. L. (2004). *ABILHAND-Kids A measure of manual ability in children with cerebral palsy*. Neurology, 63(6), 1045-1052
- Atkinson J, Hood B, Wattam-Bell J, Braddick O. *Changes in infants' ability to switch visual attention the first three months of life*. Perception. 1992;21:643–53
- Bae S, Kim KY. Dual-afferent sensory input training for voluntary movement after stroke: A pilot randomized controlled study. NeuroRehabilitation 2017; 40(3): 293-300
- Baldoli I, Cecchi F, Passetti G, Sgandurra G, Inguaggiato E, **Beani E**, Cioni G, Mihelj M, Laschi C, Dario P, *Sensorized toys for measuring the effect of home-based therapy on manipulation capabilities of infants*, submitted
- **Beani E**, Buchignani B, Cecchi F, Mannari I, Sgandurra G, Cioni G The tele-upcat project RCT: a pilot study on a sample group of children with unilateral cerebral palsy, submitted
- **Beani E,** Avola M, Sgherri G, Menici V, Sgandurra G, Cioni G Acceptability and usability of tele-UPCAT system for Action observation training in children with Unilateral Cerebral Palsy, submitted
- **Beani E**, Maselli M, Sgandurra G, Sicola E, Perazza S, Cecchi F, Dario P, Boyd R, Cioni G Asymmetry of Upper Limb activity in children with unilateral cerebral palsy: validation of a Triaxial Accelerometer Approach, submitted
- Beaton, A. V. (2003). *The nature and determinants of handedness*
- Bellelli G, Buccino G, Bernardini B, Padovani A, Trabucchi M. Action observation treatment improves recovery of postsurgical orthopedic patients: evidence for a top-down effect? Arch Phys Med Rehabil 2010; 91(10): 1489-94
- Benzies KM, Magill-Evans JE, Hayden KA, Ballantyne M. *Key components of early intervention programs for preterm infants and their parents: a systematic review and meta-analysis*. BMC Pregnancy Childbirth 2013; 13 (Suppl. 1): S10
- Berardi N, Sale A, Maffei L. *Brain structural and functional development: genetics, experience and epigenetics*. Dev Med Child Neur 2015; 57(Suppl. 2): 4–9
- Bergamini, E., Morelli, F., Marchetti, F., Vannozzi, G., Polidori, L., Paradisi, F., ... & Delussu, A. S. (2015). *Wheelchair propulsion biomechanics in junior basketball players: A method for the evaluation of the efficacy of a specific training program*. BioMed research international, 2015
- Birmingham, A. T., Wharrad, H. J., & Williams, E. J. (1985). *The variation of finger tremor with age in man*. Journal of Neurology, Neurosurgery & Psychiatry, 48(8), 788-798
- Blaga OM, Colombo J. *Visual processing and infant ocular latencies in the overlap paradigm*. Dev Psychol. 2006;42:1069–76
- Blair E, Watson L. *Epidemiology of cerebral palsy*. Semin Fetal Neonatal Med. 2006;11(2):117-125
- Bosanquet M, Copeland L, Ware R, Boyd R. *A systematic review of tests to predict cerebral palsy in young children*. Dev Med Child Neur 2013; 55: 418–26
- Braito I, Maselli M, Sgandurra G, Inguaggiato E, **Beani E**, Cecchi F, Cioni G, Boyd R, Assessment ofupper limb use in children with typical development and with neurodevelopmental disorders by inertial sensors: a systematic review, *submitted*

- Brunner IC, Skouen JS, Ersland L, Gruner R. Plasticity and response to action observation: a longitudinal FMRI study of potential mirror neurons in patients with subacute stroke. Neurorehabilitation and neural repair 2014; 28(9): 874-84
- Buccino G, Arisi D, Gough P, et al. *Improving upper limb motor functions through action observation treatment: a pilot study in children with cerebral palsy*. Developmental medicine and child neurology 2012; **54**(9): 822-8
- Buccino G, Binkofski F, Fink GR, et al. *Action observation activates premotor and parietal areas in a somatotopic manner: an fMRI study*. The European journal of neuroscience 2001; 13(2): 400-4
- Buccino G, Vogt S, Ritzl A, et al. *Neural circuits underlying imitation learning of hand actions: an event-related fMRI study.* Neuron 2004; 42(2): 323-34
- Caffarra Pea. *The effectiveness of action observation treatment (aot) in alzheimer's disease: benefit on temporal orientation and visuo-prassic abilities.* Alzheimer's & Dementia: The Journal of the Alzheimer's Association 2016; 12(7): 615
- Caligiore D, Mustile M, Spalletta G, Baldassarre G. *Action observation and motor imagery for rehabilitation in Parkinson's disease: A systematic review and an integrative hypothesis.* Neurosci Biobehav Rev 2017; 72: 210-22
- Cattaneo L CF, Jezzini A, Rizzolatti G. *Representation of goal and movements without overt motor behavior in the human motor cortex: a transcranial magnetic stimulation study*. The Journal of neuroscience the official journal of the Society for Neuroscience 2009
- Chan A-W, Tetzlaff JM, Gøtzsche PC et al. *SPIRIT 2013 Explanation and Elaboration: Guidance for protocols of clinical trials.* BMJ 2013;346:e7586.
- Chorna OD, Guzzetta A, Maitre NL, *Vision Assessments and Interventions for Infants 0-2 Years at High Risk for Cerebral Palsy: A Systematic Review*. Pediatr Neurol. 2017 Jul 20. pii: S0887-8994(17)30544-1. doi: 10.1016/j.pediatrneurol.2017.07.011. (Epub ahead of print)
- Cieza, A. Geyh, S., Chatterji, S., Kostanjsek, N., Ustund, B. & Stucki, G. (2005): *ICF linking rules: an update based on lessons learned*. J. Rehabil. Med. 37, 212–218
- Cioni G, Inguaggiato E, Sgandurra G *Early intervention in neurodevelopmental disorders: underlying neural mechanisms.* Dev Med Child Neurol. 2016 Mar;58 Suppl 4:61-6. doi: 10.1111/dmcn.13050
- Cioni G, Sgandurra G, Muzzini S, Paolicelli PB, Ferrari A. *Forms of Hemiplegia. The Spastic Forms of Cerebral Palsy: A Guide to the Assessment of Adaptive Functions*. Milano: Springer Milan; 2010: 331-56.
- Cioni, G., & Sgandurra, G. (2013). *Normal psychomotor development. Handbook of clinical neurology*, 111, 3-15
- Cobos-Carbo A, Augustovski F. *CONSORT 2010 Declaration: updated guideline for reporting parallel group randomised trials*. Med Clin (Barc) 2011;137(5):213-5. doi: 10.1016/j.medcli.2010.09.034.
- Coker-Bolt, P., Downey, R. J., Connolly, J., Hoover, R., Shelton, D., & Seo, N. J. (2017). *Exploring the feasibility and use of acceleromters before, during, and after a camp-based CIMT program for children with cerebral palsy*. Journal of Pediatric Rehabilitation Medicine, 10(1), 27-36
- Colver A, Fairhurst C, Pharoah PO. *Cerebral palsy.* Lancet 2014; 383(9924): 1240-9
- Corbetta D, Guan Y, Williams JL. *Infant eye-tracking in the context of goal-directed actions. Infancy*. 2012;17:102–25

- Cowles T, Clark A, Mares K, Peryer G, Stuck R, Pomeroy V. *Observation-to-imitate plus practice could add little to physical therapy benefits within 31 days of stroke: translational randomized controlled trial*. Neurorehabilitation and neural repair 2013; 27(2): 173-82
- Dadashi, F., Millet, G. P., & Aminian, K. (2016). *Front-crawl stroke descriptors variability assessment for skill characterisation.* Journal of sports sciences, 34(15), 1405-1412
- Database RM. *Rehab Measures: Box and Block Test*. 2012 http://www.rehabmeasures.org/Lists/RehabMeasures/DispForm.aspx?ID=917
- Database RM. *Rehab Measures: Jebsen Hand Function Test*. 2012 http://www.rehabmeasures.org/Lists/RehabMeasures/DispForm.aspx?ID=1025
- Database RM. *Rehab Measures: Quality of Upper Extremity Skills Test*. 2016
- Davila, E. M. (2011). *A comparison of bilaterally wrist-worn accelerometers on measures of free-living physical activity in adolescents* (Doctoral dissertation, Montana State University-Bozeman, College of Education, Health & Human Development)
- Davis FD. User acceptance of information technology: system characteristics, user perception and behavioral impacts. Int J Man-Machine Studies 1993; (38): 475-87
- De Felice A, Ricceri L, Venerosi A, Chiarotti F, Calamandrei G, *Multifactorial Origin of Neurodevelopmental Disorders: Approaches to Understanding Complex Etiologies.* Toxics. 2015 Mar 23;3(1):89-129. doi: 10.3390/toxics3010089.
- Deitz, J. C., Kartin, D., & Kopp, K. (2007). *Review of the Bruininks-Oseretsky test of motor proficiency, (BOT-2).* Physical & occupational therapy in pediatrics, 27(4), 87-102
- DeMatteo, C., Law, M., Russell, D., Pollock, N., Rosenbaum, P., & Walter, S. (1993*). The reliability and validity of the Quality of Upper Extremity Skills Test*. Physical & Occupational Therapy in Pediatrics, 13(2), 1-18
- Deutsch, K. M., & Newell, K. M. (2006). *Age-related changes in the frequency profile of children's finger tremor*. Neuroscience letters, 404(1), 191-195
- Donati, M., Cecchi, F., Bonaccorso, F., Branciforte, M., Dario, P., & Vitiello, N.. A modular sensorized mat for monitoring infant posture. Sensors (Basel), 2013; 14(1), 510–531
- Doyle LW, *The Victorian Infant Collaborative Study Group: Evaluation of neonatal intensive care for extremely low birthweight infants in Victoria over two decades: I. Effectiveness*. Pediatrics 2004, 113(3 Pt 1):505–509
- Dubowitz LM, Dubowitz V, Morante A. *Visual function in the newborn: a study of preterm and full-term infants*. Brain Dev. 1980;2:15–29
- Duruöz, M. T. (Ed.). (2014). *Hand function: a practical guide to assessment*. Springer Science & Business Media
- Eliasson AC1, Krumlinde-Sundholm L, Rösblad 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 Jul;48(7):549-54
- Ertelt D, Small S, Solodkin A, et al. *Action observation has a positive impact on rehabilitation of motor deficits after stroke*. *NeuroImage* 2007; 36 Suppl 2: T164-73.
- Fadiga L, Fogassi L, Pavesi G, Rizzolatti G. *Motor facilitation during action observation: a magnetic stimulation study*. *Journal of neurophysiology* 1995; 73(6): 2608-11

- Farzin F, Rivera SM, Whitney D. *Spatial resolution of conscious visual perception in infants*. Psychol Sci. 2010;21:1502–9
- Farzin F, Rivera SM, Whitney D. *Time crawls: the temporal resolution of infants' visual attention*. Psychol Sci. 2011;22:1004–10
- Floyd, A. G., Yu, Q. P., Piboolnurak, P., Wraith, E., Patterson, M. C., & Pullman, S. L. (2007). *Kinematic analysis of motor dysfunction in Niemann-Pick type C*. Clinical neurophysiology, 118(5), 1010-1018
- Franceschini M, Agosti M, Cantagallo A, Sale P, Mancuso M, Buccino G. *Mirror neurons: action observation treatment as a tool in stroke rehabilitation.* European journal of physical and rehabilitation medicine 2010; 46(4): 517-23
- Franceschini M, Ceravolo MG, Agosti M, et al. *Clinical relevance of action observation in upper-limb stroke rehabilitation: a possible role in recovery of functional dexterity. A randomized clinical trial.* Neurorehabilitation and neural repair 2012; 26(5): 456-62
- Fred D. Davis, Richard P. Bagozzi, Paul R. Warshaw, (1989) User Acceptance of Computer Technology: A Comparison of Two Theoretical Models. Management Science 35(8):982-1003
- Gatti R, Tettamanti, A., Gough, P.M., Riboldi, E., Marinoni, L., Buccino, G.,. *Action observation versus motor imagery in learning a complex motor task: a short review of literature and a kinematics study.* Neurosci 2013; Lett. 540, : 37–42
- Ghai A, Garg N, Hooda S, Gupta T. *Spasticity - Pathogenesis, prevention and treatment strategies*. Saudi journal of anaesthesia 2013; **7**(4): 453-60
- Gilmore, R., Sakzewski, L. & Boyd, R. (2009): *Upper limb activity measures for 5- to 16-year-old children with congenital hemiplegia: a systematic review*. Dev. Med. Child Neurol. 52**,** 14–21
- Gordon, A. M., Schneider, J. A., Chinnan, A., & Charles, J. R. (2007). *Efficacy of a hand–arm bimanual intensive therapy (HABIT) in children with hemiplegic cerebral palsy: a randomized control trial.* Developmental Medicine & Child Neurology, 49(11), 830-838
- Graves, L. E., Ridgers, N. D., & Stratton, G. (2008). *The contribution of upper limb and total body movement to adolescents' energy expenditure whilst playing Nintendo Wii.* European journal of applied physiology, 104(4), 617
- Gredebäck G, Johnson S, von Hofsten C. *Eye tracking in infancy research*. Dev Neuropsychol. 2010;35:1–19
- Gredebäck G, von Hofsten C. *Infants' evolving representations of object motion during occlusion: a longitudinal study of 6- to 12-month-old infants*. Infancy. 2004;6:165–84.
- Guralnick MJ. *Why early intervention works: a systems perspective.* Infants Young Child 2011; 24: 6–28
- Guzzetta A, Baldini S, Bancale A, Baroncelli L, Ciucci F, Ghirri P, Putignano E, Sale A, Viegi A, Berardi N, Boldrini A, Cioni G, Maffei L, *Massage accelerates brain development and the maturation of visual function*, The Journal of Neuroscience: the official journal of the society of neuroscience, 2009 May 6;29(18):6042-51. doi: 10.1523/JNEUROSCI.5548-08.2009
- Guzzetta A, Cioni G, Cowan F, Mercuri E. *Visual disorders in children with brain lesions: 1. Maturation of visual function in infants with neonatal brain lesions: correlation with neuroimaging*. Eur J Paediatr Neurol. 2001;5:107–14
- Hack M, Flannery DJ, Schluchter M, Cartar L, Borawski E, Klein N, *Outcomes in young adults for very-low-birthweight infants*. N Engl J Med 2002, 346(3):149–157
- Hadders-Algra M. *Early diagnosis and early intervention in cerebral palsy*. Front Neurol 2014; 5: 185

- Hanna SE, Russell DJ, Bartlett DJ, Kertoy M, Rosenbaum PL, Wynn K. *Measurement practices in pediatric rehabilitation: a survey of physical therapists, occupational therapists, and speech-language pathologists in Ontario* Phys Occup Ther Pediatr. 2007;27(2):25-42
- Hanna, S., Russell, D. Bartlett D., Kertoy, M., Rosenbaum, P. & Swinton, M. (2005): *Clinical measurement: Practical guidelines for service providers*. CanChild Centre for Childhood disability research. McMaster University
- Heckman JJ. *Schools, skills, and synapses*. Econ Inq. 2008;46:289
- Hoffmann TC, Glasziou PP, Boutron I et al. *Better reporting of interventions: template for intervention description and replication (TIDieR) checklist and guide*. BMJ 2014; 348:g1687. doi: 10.1136/bmj.g1687.
- Holmefur M, Aarts P, Hoare B, Krumlinde-Sundholm L *Test-retest and alternate forms reliability of the assisting hand assessment*. J Rehabil Med. 2009 Nov;41(11):886-91. doi: 10.2340/16501977-0448
- Holmefur M, Krumlinde-Sundholm L, Eliasson AC. *Interrater and intrarater reliability of the Assisting Hand Assessment*. The American journal of occupational therapy : official publication of the American Occupational Therapy Association 2007; 61(1): 79-84
- House JH, Gwathmey FW, Fidler MO. *A dynamic approach to the thumbin- palm deformity in cerebral palsy: evaluation and results in fifty-six patients*. J Bone Joint Surg 1981;63:216-25.
- Inguaggiato E, Sgandurra G, Bargagna S, **Beani E**, Orlando M, Purpura G, Cecchi F, Dario P, Cioni G and CareToy Consortium *CareToy Early Intervention in infants with Down Syndrome: a feasibility study*, submitted
- Inguaggiato, E., Sgandurra, G., **Beani, E**., Piccardo, G., Giampietri, M., Bartalena, L., & Cioni, G. *CareToy project for early intervention in infants born preterm: Preliminary findings on parent-infant interaction*. Giornale di Neuropsichiatria dell'Eta` Evolutiva, 2015; 35, 147–154 (in Italian Language)
- Jebsen, R. H. (1969*). An objective and standardized test of hand function*. Arch. Phys. Med. Rehabil., 50, 311-319
- Johnson SP, Amso D, Slemmer JA. *Development of object concepts in infancy: evidence for early learning in an eyetracking paradigm*. Proc Natl Acad Sci. 2003;100:10568–73.
- Johnson SP, Slemmer JA, Amso D. *Where infants look determines how they see: eye movements and object perception performance in 3-month-olds*. Infancy. 2004;6:185–201
- Johnston MV, *Clinical disorders of brain plasticity*, Brain and Development, vol. 26, no. 2, pp. 73–80, 2004
- Johnston MV, Ishida A, Ishida WN, Matsushita HB, Nishimura A, Tsuji M. *Plasticity and injury in the developing brain*. Brain Dev. 2009;31:1- 10. 6
- Kaneko, M., Yamashita, Y., & Iramina, K. (2016). *Quantitative Evaluation System of Soft Neurological Signs for Children with Attention Deficit Hyperactivity Disorder*. Sensors, 16(1), 116
- Kaneko, M., Yamashita, Y., Inomoto, O., & Iramina, K. (2015). *Soft neurological signs in childhood by measurement of arm movements using acceleration and angular velocity sensors*. Sensors, 15(10), 25793-25808
- Kim CH, Bang DH. *Action observation training enhanced upper extremity function in sub acute stroke survivor with moderate impairment: a double-blind, randomized controlled pilot trial*. Journal of the korean society of physical medicine, 2016
- Kim E, Kim K. *Effect of purposeful action observation on upper extremity function in stroke patients*. Journal of physical therapy science 2015; 27(9): 2867-9.

- Kim E, Kim K. *Effects of purposeful action observation on kinematic patterns of upper extremity in individuals with hemiplegia*. Journal of physical therapy science 2015; 27(6): 1809-11
- Kim JC, Lee HM. *The Effect of Action Observation Training on Balance and Sit to Walk in Chronic Stroke: A Crossover Randomized Controlled Trial*. Journal of motor behavior 2017: 1-8
- Kim JH, Chung EJ, Lee BH. A *study of analysis of the brain wave with respected to action observation and motor imagery: a pilot randomized controlled trial*. Journal of physical therapy science 2013; 25(7): 779-82
- Kim JH, Lee BH. *Action observation training for functional activities after stroke: a pilot randomized controlled trial*. NeuroRehabilitation 2013; 33(4): 565-74
- Kim JY, Kim JM, Ko EY. *The effect of the action observation physical training on the upper extremity function in children with cerebral palsy.* Journal of exercise rehabilitation 2014; 10(3): 176-83
- Kirkpatrick E, Pearse J, James P, Basu A. *Effect of parent-delivered action observation therapy on upper limb function in unilateral cerebral palsy: a randomized controlled trial* Developmental medicine and child neurology 2016; 58(10): 1049-56
- 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*. Clinical rehabilitation 2010; 24(10): 887-900
- Koman LA, Williams RM, Evans PJ, et al. *Quantification of upper extremity function and range of motion in children with cerebral palsy*. Developmental medicine and child neurology 2008; 50(12): 910-7
- Krägeloh-Mann I & Cans C. *Cerebral palsy update*. Brain Dev 2009;7:537-44
- Krumlinde-Sundholm L. *Choosing and using assessments of hand function*. In: Eliasson AC, Burtner PA, editors. *Improving hand function in children with cerebral palsy: theory, evidence and intervention*. London: Mac Keith Press, 2008: 176–97
- Krumlinde-Sundholm, L., Holmefur, M., Kottorp, A., & Eliasson, A. C. (2007). *The Assisting Hand Assessment: current evidence of validity, reliability, and responsiveness to change*. Developmental Medicine & Child Neurology, 49(4), 259-264
- Law M, MacDermid JC. *Evidence-based rehabilitation: a guide to practice*. 2nd edition. Philadelphia: Slack Publishing; 2008
- Le Moing, A. G., Seferian, A. M., Moraux, A., Annoussamy, M., Dorveaux, E., Gasnier, E., ... & Servais, L. (2016). *A Movement Monitor Based on Magneto-Inertial Sensors for Non-Ambulant Patients with Duchenne Muscular Dystrophy: A Pilot Study in Controlled Environment.* PloS one, 11(6), e0156696
- Lee D, Roh H, Park J, Lee S, Han S. *Drinking behavior training for stroke patients using action observation and practice of upper limb function.* Journal of physical therapy science 2013; 25(5): 611-4
- Lee HJ, Kim YM, Lee DK. *The effects of action observation training and mirror therapy on gait and balance in stroke patients.* Journal of physical therapy science 2017; 29(3): 523-6
- Lemmens, R. J., Seelen, H. A., Timmermans, A. A., Schnackers, M. L., Eerden, A., Smeets, R. J., & Janssen-Potten, Y. J. (2015). *To What Extent Can Arm–Hand Skill Performance—of Both Healthy Adults and Children—Be Recorded Reliably Using Multiple Bodily Worn Sensor Devices*?. IEEE Transactions on Neural Systems and Rehabilitation Engineering, 23(4), 581-590
- Löwing K, Bexelius A, Carlberg EB*: Goal-directed functional therapy: a longitudinal study on gross motor function in children with cerebral palsy*. Disabil Rehabil 2010, 32(11):908–916
- Lundqvist-Persson C, Lau G, Nordin P, Bona E, Sabel KG, *Preterm infants' early developmental status is associated with later developmental outcome*. Acta Paediatr 2012, 101(2):172–178

- MacArthur, B., Coe, D., Sweet, A., & Raynor, H. (2014). *Active videogaming compared to unstructured, outdoor play in young children: percent time in moderate-to vigorous-intensity physical activity and estimated energy expenditure.* Games for health journal, 3(6), 388-394
- MacDermid JC1, Grewal R, MacIntyre NJ, *Using an evidence-based approach to measure outcomes in clinical practice*, Hand Clin. 2009 Feb;25(1):97-111, vii. doi: 10.1016/j.hcl.2008.11.001
- Mackintosh, K. A., Montoye, A. H. K., Pfeiffer, K. A., & McNarry, M. A. (2016). *Investigating optimal accelerometer placement for energy expenditure prediction in children using a machine learning approach*. Physiological measurement, 37(10), 1728
- Manetti S, Cecchi F, Sgandurra G, Cioni G, Laschi C, Dario P, Turchetti G. *Early Stage Economic Evaluation of Caretoy System for Early Intervention In Preterm Infants At Risk of Neurodevelopmental Disorders.* Value Health. 2015; 18(7):A358
- Manual Ability Classification System. *Manual Ability Classification System for children with cerebral palsy 4-18 years*. http://www.macs.nu (accessed 09.2017)
- Mathiowetz V FS, Wiemer D. *Box and block test of manual dexterity; norms for 6–19 year olds.* Canadial Journal of Occupational Therapy 1985; **52**: 241-5
- Meyer, M. R. U., Baller, S. L., Mitchell, S. M., & Trost, S. G. (2013). *Comparison of 3 accelerometer data reduction approaches, step counts, and 2 self-report measures for estimating physical activity in free-living adults*. Journal of Physical Activity and Health, 10(7), 1068-1074.
- Morante A, Dubowitz LM, Leven M, Dubowitz V, *The development of visual function in normal and neurologically abnormal preterm and fullterm infants*. Dev Med Child Neurol. 1982;24:771–84
- Morgan C, Novak I, Badawi N, *Enriched Environments and Motor Outcomes in Cerebral Palsy: systematic review and meta-analysis*. Pediatrics 2013, 132(3):e735–e746
- Morris S. *Ashworth and Tardieu scales: their clinical relevance for measuring spasticity in adult and paediatric neurological populations*. Phys Ther Rev 2002; **7**: 53-72
- Mukhopadhyay, S. C. (2015). *Wearable sensors for human activity monitoring: A review*. IEEE sensors journal, 15(3), 1321-1330
- Mulder T. *Motor imagery and action observation: cognitive tools for rehabilitation.* Journal of neural transmission 2007; 114(10): 1265-78
- Mutlu A, Livanelioglu A, Gunel MK. *Reliability of Ashworth and Modified Ashworth scales in children with spastic cerebral palsy*. BMC musculoskeletal disorders 2008; 9: 44
- Neil PA, Chee-Ruiter C, Scheier C, Lewkowicz DJ, Shimojo S. *Development of multisensory spatial integration and perception in humans*. Dev Sci. 2006;9:454–64
- Noorkõiv, M., Rodgers, H., & Price, C. I. (2014). *Accelerometer measurement of upper extremity movement after stroke: a systematic review of clinical studies*. Journal of neuroengineering and rehabilitation, 11(1), 144
- Novak I, Cusick A, Lowe K. A pilot study on the impact of occupational therapy home programming for young children with cerebral palsy. The American journal of occupational therapy : official publication of the American Occupational Therapy Association 2007; 61(4): 463-8
- Novak I, Morgan C, Adde L, Blackman J, Boyd RN, Brunstrom-Hernandez J, Cioni G, Damiano D, Darrah J, Eliasson AC, et al, *Early, Accurate Diagnosis and Early Intervention in Cerebral Palsy: Advances in Diagnosis and Treatment.* JAMA Pediatr. 2017 Sep 1;171(9):897-907. doi: 10.1001/jamapediatrics.2017.1689.

- O'Neil, M. E., Fragala-Pinkham, M., Lennon, N., George, A., Forman, J., & Trost, S. G. (2016). *Reliability and validity of objective measures of physical activity in youth with cerebral palsy who are ambulatory*. Physical therapy, 96(1), 37-45
- Oskoui M, Coutinho F, Dykeman J, Jetté N, Pringsheim T. *An update on the prevalence of cerebral palsy: a systematic review and meta-analysis*. Dev Med Child Neurol. 2013;55(6): 509-519
- Park EC, Hwangbo G. *The effects of action observation gait training on the static balance and walking ability of stroke patients*. Journal of physical therapy science 2015; 27(2): 341-4
- Park HJ, Oh DW, Choi JD, et al. *Action observation training of community ambulation for improving walking ability of patients with post-stroke hemiparesis: a randomized controlled pilot trial*. Clinical rehabilitation 2017; 31(8): 1078-86
- Park HR, Kim JM, Lee MK, Oh DW. *Clinical feasibility of action observation training for walking function of patients with post-stroke hemiparesis: a randomized controlled trial.* Clinical rehabilitation 2014; 28(8): 794-803
- Passetti G, Cecchi F, Baldoli I, Sgandurra G, **Beani E**, Cioni G, Laschi C, Dario P *Sensorized toys for measuring manipulation capabilities of infants at home* Conf Proc IEEE Eng Med Biol Soc. 2015;2015:7390-3. doi: 10.1109/EMBC.2015.7320099
- Pomeroy VM, Clark CA, Miller JS, Baron JC, Markus HS, Tallis RC. *The potential for utilizing the "mirror neurone system" to enhance recovery of the severely affected upper limb early after stroke: a review and hypothesis*. Neurorehabilitation and neural repair 2005; 19(1): 4-13
- Pratesi A, Cecchi F, **Beani E**, Sgandurra G, Cioni G, Laschi C, Dario P, *A new system for quantitative evaluation of infant gaze capabilities in a wide visual field* Biomed Eng Online. 2015 Sep 7;14:83. doi: 10.1186/s12938-015-0076-7
- Quinn PC, Doran MM, Reiss JE, Hoffman JE. *Time course of visual attention in infant categorization of cats versus dogs: evidence for a head bias as revealed through eye tracking*. Child Dev. 2009;80:151–61
- R. Chen, L. G. Cohen, and M. Hallett, *Nervous system reorganization following injury*, Neuroscience, vol. 111, no. 4, pp. 761–773, 2002
- Randall M, Carlin JB, Chondros P, Reddihough D. *Reliability of the Melbourne assessment of unilateral upper limb function.* Developmental medicine and child neurology 2001; **43**(11): 761-7
- Rehab-Scales. *ABILHAND-Kids*
- Riethmuller AM, Jones R, Okely AD, *Efficacy of interventions to improve motor development in young children: a systematic review*. Pediatrics 2009, 124(4):e782–e792
- Rihar A, Sgandurra G, **Beani E**, Cecchi F, Pašič J, Cioni G, Dario P, Mihelj M, Munih M *CareToy: Stimulation and Assessment of Preterm Infant's Activity Using a Novel Sensorized System.* Ann Biomed Eng. 2016 Dec;44(12):3593-3605. Epub 2016 Jun 10.
- Rihar, A., Mihelj, M., Pasˇic ˇ, J., Kolar, J., & Munih, M. *Infant trunk posture and arm movement assessment using pressure mattress, inertial and magnetic measurement units (IMUs)*. Journal of NeuroEngineering and Rehabilitation,2014; 11(1), 133
- Rizzolatti G, Fadiga L, Matelli M, et al. *Localization of grasp representations in humans by PET: 1. Observation versus execution*. Experimental brain research 1996; 111(2): 246-52
- Rizzolatti G, Sinigaglia C. *The mirror mechanism: a basic principle of brain function*. Nature reviews Neuroscience 2016; 17(12): 757-65

- Rosenbaum, P. & Stewart, D. (2004*): The World Health Organization International Classification of Functioning,Disability, and Health: a model to guide clinical thinking, practice and research in the field of cerebral palsy*. Semin. Pediatr. Neurol. 11, 5–10
- Rosenbaum, P. (1998). *Screening tests and standardized assessments used to identify and characterize developmental delays.* Semin. Pediatr. Neurol. **5,** 27–32
- Sakzewski L, Ziviani J, Boyd R*: Systematic Review and Meta-analysis of Therapeutic Management of Upper-Limb Dysfunction in Children with congenital hemiplegia*. Pediatrics 2009, 123:e1111-1122
- Sale A, Berardi N, Maffei L, *Enrich the environment to empower the brain* Trends Neurosci 2009, 32(4):233–239
- Sale A, Berardi N, Maffei L. *Environment and brain plasticity: towards an endogenous pharmacotherapy*. Physiol Rev 2014; 94: 189–234.
- Sale P, Ceravolo MG, Franceschini M. *Action observation therapy in the subacute phase promotes dexterity recovery in right-hemisphere stroke patients*. BioMed research international 2014; 2014: 457538
- Sgandurra G, Bartalena L, Cecchi F, et al. *A pilot study on early home-based intervention through an intelligent baby gym (CareToy) in preterm infants*. Res Dev Disabil, 2016; 53-54:32-42
- Sgandurra G, Buchignani B, Iacono O, Inguaggiato E, **Beani E**, Sicola E, Feys H, Cioni G, *Action observation for upper and lower limb rehabilitation in hemiplegia: a systematic review*, submitted
- Sgandurra G, Cecchi F, **Beani E**, Mannari I, Maselli M, Falotico FP, Inguaggiato E, Perazza S, SicolaE, Feys H, Klingels K, Ferrari A, Dario P, Boyd R, Cioni G *Tele-UPCAT: Study Protocol of a Randomized Controlled Trial of a Home-Based Tele-monitored UPper Limb Children Action Observation Training for Participants with Unilateral Cerebral Palsy*, submitted
- Sgandurra G, Cioni G, *Clinical and instrumental assessment of the upper limb in cerebral palsy* in Mercuri, E. Fedrizzi and G. Cioni *New Diagnostic and Therapeutic Tools in Paediatric Neurology*, E. 2011 John Libbey Eurotext, pp. 119–132.
- Sgandurra G, Ferrari A, Cossu G, Guzzetta A, Biagi L, RìTosetti M, Fogassi L, Cioni G, *Upper limb children action-observation training (UP-CAT): a randomised controlled trial in Hemiplegic Cerebral Palsy*, BMC Neurology 2011 11:80
- Sgandurra G, Ferrari A, Cossu G, Guzzetta A, Fogassi L, Cioni G. *Randomized trial of observation and execution of upper extremity actions versus action alone in children with unilateral cerebral palsy*. Neurorehabil Neural Repair. 2013 Nov-Dec;27(9):808-15. doi: 10.1177/1545968313497101. Epub 2013 Jul 25.
- Sgandurra G, Lorentzen J, Inguaggiato E, Bartalena L, **Beani E**, Cecchi F, Dario P, Giampietri M, Greisen G, Herskind A, Nielsen JB, Rossi G, Cioni G; CareToy Consortium. *A randomized clinical trial in preterm infants on the effects of a home-based early intervention with the 'CareToy System* PLoS One. 2017 Mar 22;12(3):e0173521 doi:10.1371/journal.pone.0173521. eCollection 2017
- Siegert RJ, Taylor WJ*: Theoretical aspects of goal-setting and motivation in rehabilitation*. Disabil Rehabil 2004, 26(1):1–8
- Simmonds MJ, 1997. *Muscle strength* in Van Deusen J, Brunt D, editors, *Assessment in Occupational Therapy and Physical therapy*. Philadelphia WB Saunders pp 27-48
- Spittle A, Orton J, Anderson P, Boyd R, Doyle LW, *Early developmental intervention programmes post-hospital discharge to prevent motor and cognitive impairments in preterm infants*. Cochrane Database Syst Rev 2012, 12:CD005495

- Spittle A, Orton J, Doyle LW, Boyd R, *Early developmental intervention programs post hospital discharge to prevent motor and cognitive impairments in preterm infants (Review)*, The Cochrane Database of Systematic Reviews. 2007 Apr 18;(2):CD005495
- Stavsky M, Mor O, Mastrolia SA, Greenbaum S, Than NG, Erez O. *Cerebral Palsy-Trends in Epidemiology and Recent Development in Prenatal Mechanisms of Disease, Treatment, and Prevention*. Frontiers in pediatrics 2017; 5: 21
- Sugg K, Muller S, Winstein C, Hathorn D, Dempsey A. *Does Action Observation Training With Immediate Physical Practice Improve Hemiparetic Upper-Limb Function in Chronic Stroke?* Neurorehabilitation and neural repair 2015; 29(9): 807-17
- Vohr BR, Wright LL, Poole WK, McDonald SA. *Neurodevelopmental outcomes of extremely low birthweight infants <32 weeks' gestation between 1993 and 1998*. Pediatrics 2005; 116: 635–43
- Wagner LV, Davids JR. *Assessment tools and classification systems used for the upper extremity in children with cerebral palsy* Clin Orthop Relat Res. 2012 May;470(5):1257-71. doi: 10.1007/s11999-011-2065-x
- World Health Organization (WHO). (2001): *The International Classification of Functioning, Disability and Health (ICF)*. Geneva: World Health Organization, 2001
- World Health Organization, World Bank. *World Report on Disability*. Geneva: World Health Organization, 2011 (www.who.int/disabilities/world_report/2011/en/index.html
- World Health Organization. (2007): *International Classification of Functioning, Disability and Health Children & Youth Version (ICF-CY)*. Geneva: World Health Organization
- Zhu MH, Wang J, Gu XD, et al. *Effect of action observation therapy on daily activities and motor recovery in stroke patients*. International Journal of Nursing Sciences 2015; 2(3): 279-82