國立臺灣大學醫學院物理治療系暨研究所

博士論文

School and Graduate Institute of Physical Therapy

College of Medicine

National Taiwan University

Doctoral Dissertation

透過行動應用程式進行父母感知與人工智慧嬰兒動作評估的效度研究

A Validity Study of Parental Perception and Artificial Intelligence of Infant Motor Assessment via Mobile Application

Yohanes Purwanto

指導教授:鄭素芳 博士

Advisor: Suh-Fang Jeng, PT, ScD

中華民國 114 年 9 月

September, 2025

doi:10.6342/NTU202504540

國立臺灣大學醫學院物理治療系暨研究所

博士論文

School and Graduate Institute of Physical Therapy
College of Medicine
National Taiwan University

Doctoral Dissertation

透過行動應用程式進行父母感知與人工智慧嬰兒動作評估的效度研究

A Validity Study of Parental Perception and Artificial Intelligence of Infant Motor Assessment via Mobile Application

Yohanes Purwanto

指導教授:鄭素芳 博士

Advisor: Suh-Fang Jeng, PT, ScD

中華民國 114 年 9 月

September, 2025

doi:10.6342/NTU202504540

DOCTORAL DISSERTATION ACCEPTANCE CERTIFICATE

NATIONAL TAIWAN UNIVERSITY

A Validity Study of Parental Perception and Artificial Intelligence

of Infant Motor Assessment via Mobile Application

The undersigned, appointed by the School and Graduate Institute of Physical Therapy on September 22, 2025, have examined a Doctoral Dissertation entitled above, presented by Yohanes Purwanto, D10428003, candidate, and hereby certify that it is worthy of acceptance

Oral examination committee:

Suh-Fang Jeng, PT, ScD **(Advisor)**	Jane Yung-jen Hsu, MS, PhD	Po-Nien Tsao, MD, PhD

Jenq Lin, PT, PhD **(Director)**

doi:10.6342/NTU202504540

Acknowledgment

First of all, I am deeply grateful to my advisor, Dr. Suh-Fang Jeng, for her patient guidance in shaping me not only as a practitioner but also by instilling in me the philosophy of research and academia, which will be invaluable for the development of my professional career in Indonesia. I sincerely thank my advisor, Dr. Suh-Fang Jeng, and committee member, Dr. Jane Yung-Jen Hsu, for their invaluable guidance throughout this research. I am also grateful to the interdisciplinary team members who facilitated this study, including Dr. Po-Nien Tsao, Dr. Ting-An Yen, Dr. Wei-Chih Liao, and Dr. Wei J. Chen.

I would like to acknowledge the infants and their families for participating in this study, as well as the medical, nursing, and physical therapy staff at National Taiwan University Children's Hospital for their assistance with recruitment. My appreciation also goes to all members of the iAgents Lab in the Department of Computer Science & Information Engineering at NTU for their contributions, especially Emily Chen, Erick Chandra, Chin-Yi Liao, Yi-Ning Chen, Jian-Chi Chen, Wei Jie Chen, Ya-Yun Cheng, and Zi-Xiang Wei. I also wish to thank Shiang-Chin Lin, Yu-Ching Hsiao, Chun-Wun Hsieh, and Qian-Wei Chen of the Infant Motor Development Laboratory at the School and Graduate Institute of Physical Therapy at NTU for their support with developmental assessments and follow-up. Finally, I would like to express my gratitude to my family and colleagues in Indonesia who have supported me throughout my doctoral studies.

doi:10.6342/NTU202504540

Abstract

Background: Preterm birth is a global healthcare issue because of the increased risk of neurodevelopmental impairments, with motor disorders being among the most common. Early motor assessment of preterm infants is crucial to help identify those who may develop motor disorders and require early intervention. Mobile applications integrated with artificial intelligence (AI) technology for infant motor assessment are increasingly used. However, such innovations have rarely been designed for infants older than three months. **Purpose:** This study had three purposes: (1) to upgrade a mobile application *"Baby Go"* to contain a parental perception assessment and an AI model for infant motor assessment in full-term and preterm infants, (2) to examine the agreement between the parental perception and AI results from the application, and (3) to evaluate the agreement of the parental perception and the AI results with physiotherapists' labelling results. **Methods:** Preterm and full-term infants were recruited from the National Taiwan University Children's Hospital in Taiwan. The *"Baby Go"* application was upgraded to contain the features of parental perception of infant movements, AI assessment for 38 movements, developmental follow-up, and parental education. Parents were asked to register the *"Baby Go"* application and use it regularly. The validity of the parental perception and the AI assessment results was examined using physiotherapists' labelling results as the gold standard. **Results:** Fifty parents registered the application, of which 45 parents (29 of preterm infants and 16 of full-term infants) uploaded at least one video. Movement-based assessment on 498 videos without technical problems showed that the agreement between the parents' and physiotherapists' results was 79%, the agreement between the

doi:10.6342/NTU202504540

physiotherapists' and AI results was 79%, and the agreement between the parents' and AI results was 73%. Age-based assessment on 61 completed trials without technical problems showed that the agreement between the parents' and AI results was 97%, the agreement between the physiotherapists' and AI results was 82%, and the agreement between the parents' and physiotherapists' results was 74%. Concurrent validity of the parental perception and AI results on 14 trials compared with the AIMS onsite assessment results showed agreement of 79%, sensitivity of 100%, specificity of 75%, positive predictive value of 40%, and negative predictive value of 100%. **Conclusions:** The upgraded *'Baby Go'* application successfully integrates AI-based infant motor assessment but requires further development of the parental perception feature. AI assessments show nearly high agreement for movement-based assessment and high agreement for age-based evaluations with the videos with technical problems excluded. Future work needs to mitigate technical issues and may consider re-arrangement of movement sets to enhance the AI performance for remote infant motor assessment. The *"Baby Go"* application can serve as a supplement for parents in early infant motor screening and for pediatricians and physiotherapists in motor development follow-up and early intervention, but its parental observations and self-assessment have always followed the standardized assessments designed and monitored by health professionals.

Keywords: preterm infants, mobile application, artificial intelligence, infant motor assessment.

doi:10.6342/NTU202504540

Table of Contents

doi:10.6342/NTU202504540

doi:10.6342/NTU202504540

doi:10.6342/NTU202504540

Chapter 1. Literature Review

1.1 Preterm Infants and Neurodevelopmental Outcomes

Preterm birth is a worldwide epidemic, with the global estimation of preterm birth as approximately 15 million per year (Blencowe et al., 2013). Preterm birth is defined as infant born before 37 weeks of gestation age (Cao et al., 2022). At the global level, there has been no change in the rate of preterm birth among live births in the recent decade (9.8% in 2010 and 9.9% in 2020) (Ohuma et al., 2023). However, the rate of preterm birth in southern Asia has been higher (13.2%) than that of other regions with higher incidences in low- and middle-income nations, including Bangladesh (16.2%), Pakistan (14.4%), and India (13.0%) (Ohuma et al., 2023). The persistent high rate of preterm birth in the past decade warrants the need to assess the outcome in those surviving preterm infants.

Preterm birth is a critical global problem because of its implications for morbidity and mortality (Chawanpaiboon et al., 2019). Advances in medical technologies and therapeutic perinatal and neonatal care have substantially improved the rates of survival among preterm infants, even at the lowest gestational ages (Lima & Mondardini, 2023). However, compared to those infants born full-term, the survival of preterm neonates is burdened by increased rates of short- and long-term morbidities, ranging from major disabilities such as cerebral palsy, mental retardation, and sensory impairments to minor disabilities like language and learning problems, attention-deficit hyperactivity disorder, behavioral difficulties, and neuromotor delays (Frey & Klebanoff, 2016).

Our research group has extensively examined the developmental outcomes of preterm infants and full-term infants in Taiwan. Early studies demonstrated that very low birth weight

doi:10.6342/NTU202504540

(VLBW) preterm infants exhibit distinct neurobehavioral profiles at term age (Jeng et al., 1998). Another study showed that preterm infants without major neonatal disease attain independent walking at a significantly later age than term infants, with shorter stride lengths at 18 months of corrected age (Jeng et al., 2008). Further research found that children born with VLBW have significantly lower cognitive performance at primary school age compared with full-term peers (Mu et al., 2008). In-hospital developmental care has short-term benefits for VLBW preterm infants in reducing the risk of retinopathy and feeding desaturation, as well as in enhancing weight gains at term age (Chen et al., 2013). Moreover, family-centered care and environmental factors were shown to influence developmental outcomes, highlighting the importance of early assessment, intervention, and supportive care in optimizing long-term neurodevelopment (Yu et al., 2017). These studies underscore the multifactorial influences on preterm infant development and the value of early monitoring and intervention.

Attainment of motor skills in an infant affords opportunities to explore the environment to develop perceptual, motor, cognitive, social, and emotional behavior (Rosenbaum et al., 2001). A longitudinal study of healthy preterm infants and full-term infants aged 4 to 48 months showed a significant association between early motor development and later cognitive function (Piek et al., 2008). Research evidence has demonstrated that motor skills have a critical impact on other developmental domains, including cognitive, social, and emotional development (Su et al., 2017; Zhang et al., 2023). Preterm infants with very low birth weight demonstrate quantitative and qualitative differences in motor development compared to term infants during the first years of life. Specifically, three distinct motor trajectories have been identified: stably normal (55%), deteriorating (32%), and persistently delayed (13%), and these trajectories during the first year of life were predictive of motor and mental outcomes at 24 months (Su et al., 2017). These findings

doi:10.6342/NTU202504540

suggest that early motor assessment is valuable to detect signs of neurodevelopmental disorders in preterm infants (Baccinelli et al., 2020).

1.2 Infant Motor Assessment

The American Academy of Pediatrics (AAP) recommends developmental surveillance at all preventive care visits and standardized developmental screening of children at ages 9, 18, 30, and 48 months (Noritz et al., 2013). Employment of evidence-based standardized developmental screening tools helps identify those infants who may have developmental disorders and make timely referrals for early diagnostic assessment and early intervention (Agarwal et al., 2020).

Infant motor assessment tools consist of screening tests and diagnostic assessments. Diagnostic tools typically administered by medical professionals using standardized tools and equipment are usually available at hospitals for making developmental diagnoses (Mendonça et al., 2016). The purpose of screening is to identify children who appear to have delays in development and to make an appropriate referral for diagnostic assessment. Screening tools are often based on observations by parents or caregivers, and may involve simple questionnaires or checklists that are relatively short and inexpensive (Tieman et al., 2005).

A good screening tool should be inexpensive, simple to administer, have good psychometric properties, be culturally appropriate, and require minimal training (Mendonça et al., 2016). Developmental screening tools administered by a trained professional are exemplified by Denver Developmental Screening Tool-II (DDST-II) with a high specificity of 83% and low sensitivity of 43% (Hansen, 2021), the Bayley Neurodevelopmental screener with a low sensitivity of 21 – 40 % and a high specificity of 66 - 78% (Hess et al., 2004), the Parents' Evaluation of Developmental Status (PEDS) with a high sensitivity of 83% and a high specificity of 84%, and the Ages and Stages Questionnaires (ASQ-3) with a moderate sensitivity of 56.3%

doi:10.6342/NTU202504540

and a high specificity of 93.0% (Agarwal et al., 2020). Various developmental screening tools exhibit different levels of sensitivity and specificity.

When developmental screening identifies infants as being at risk for a developmental disorder, subsequent standardized diagnostic assessment should be pursued and will be conducted by healthcare professionals. The results of diagnostic assessments provide the information for developmental diagnosis, allowing prompt initiation of specific and appropriate early childhood therapeutic interventions (Lipkin et al., 2020). Several standardized assessment tools have been modified to optimize the availability of professional resources, the instrument's cost, the suitability of the results, and the administration time (Gücüyener et al., 2006; Suir et al., 2022). For example, the Bayley Infant Neurodevelopmental Screening (BINS) was modified from items of the Bayley Scales of Infant Development (BSID-II), which demonstrated moderate to high test-retest reliability and internal consistency (Aylward, 1995). The TIMPSI, a short version of the Test of Infant Motor Performance (TIMP), is used for screening. which showed high accuracy but low negative and positive predictive values (Campbell et al., 2008). Although these modified infant motor screening tests have simplified the testing procedure and enhanced accessibility, improvement of their validity is warranted to better detect infants who have motor disorders.

Healthcare professionals and researchers require standardized diagnostic assessment tools to identify and classify motor problems in children. The diagnostic assessments are such as the Bayley Scale of Infant and Toddler Development-III (Bayley-III) (Bayley, 2006), Bruininks-Oseretsky Test of Motor Proficiency-2 (BOT-2) (Bruininks & Bruininks, 1978), Movement Assessment Battery for Children-2 (MABC-2) (Henderson et al., 1992), McCarron Assessment of Neuromuscular Development (MAND) (McCarron, 1982), Neurological Sensory Motor

doi:10.6342/NTU202504540

Developmental Assessment (NSMDA) (Burns, 1992), Peabody Developmental Motor Scales-2 (PDMS-2) (Folio, 1983), Test of Gross Motor Development-2 (TGMD-2) (Ulrich, 1985), Alberta Infant Motor Scale (AIMS) and General Movement Assessment (GMA) (Piper & Darrah, 2021).

Psychometric data of the standardized diagnostic assessment tools varied from moderate to excellent values (Griffiths et al., 2018). Validity and internal consistency varied from fair to excellent (α=0.5–0.99) (Griffiths et al., 2018). The Bayley-III motor scale showed high intra-rater reliability (intraclass correlation coefficient ICC) = 0.85–0.98, the NSMDA showed a high positive predictive value of 0.85, and the MABC-2 showed a high negative predictive value of 0.97. Test–retest reliability is excellent in the BOT-2 (ICC=0.80–0.99), the PDMS-2 (ICC=0.97), MABC-2 (ICC=0.83–0.96), and the TGMD-2 (ICC=0.81–0.92). TGMD-2 has high inter-rater (ICC=0.88–0.93) and intra-rater reliability (ICC=0.92–0.99) (Griffiths et al., 2018), the AIMS has high inter-rater (ICC=0.99) and intra-rater reliability (ICC=0.97), the GMA Global score (ICC=0.95) and motor optimality score have high reliability (ICC=0.90) (Grant et al., 2023). These assessments play a crucial role in diagnosing and evaluating infant motor development. It is essential to consider the psychometric data about reliability and validity and specific strengths and limitations of each diagnostic assessment tool when selecting the appropriate assessment for a particular case.

1.3 Home-based Infant Motor Assessment

Infants should receive periodic developmental screening using a standardized test, as recommended in the AAP policy statement, to identify infants with developmental disorders in the medical home (Lipkin et al., 2020). Infant motor assessment may be assisted by parents at home (Noritz et al., 2013). It is crucial for professionals to emphasize the importance of early

doi:10.6342/NTU202504540

assessment and provide educational resources to parents, fostering a collaborative approach in supporting infants with identified developmental concerns.

Cultural studies provide evidence to support the relationship between the home environment and infant motor development (Abbott et al., 2000). The developmental studies in South America (Brazil), Australia, Africa (Cameroon, Malawi), North America (USA), European (Belgium, Germany, Greece, Denmark, Netherlands), and Asia (Cambodia, India, Singapore, Sri Lanka, Taiwan, Thailand) suggest that the timing of motor development varied among infants growing up in different regions (Mendonça et al., 2016). The gross motor acceleration observed in these populations has been attributed to culture-specific caregiving practices, including formal handling routines, positioning, and cultural beliefs and expectations (Mendonça et al., 2016). For example, practicing standing and sitting and applying massage or stretching of the limbs is common in African and Caribbean cultures but less common in Western cultures (Oudgenoeg-Paz et al., 2020). Delayed onset of gross motor milestones has been documented in Chinese and Japanese infants, which may relate to less vigorous handling and highly protective and avoidance of prone positioning compared to those of Aboriginal and African mothers (Mendonça et al., 2016; Wang, 2022). Healthcare professionals should consider cultural differences in parental expectations and caregiving that may impact the timing and progression of motor milestones.

The use of home videos made by parents to assess infant development has been the subject in recent research (Boonzaaijer et al., 2019). Home videos have been used as a valuable resource for assessment in pediatric rehabilitation (Yeh et al., 2020) because of their feasibility and natural environmental context (Fish & Jones, 2021). In addition, home videos may reduce the overall burden of current clinical testing procedures on infants and parents (Boonzaaijer et

doi:10.6342/NTU202504540

al., 2017). Home video assessments made by parents provide a feasible and informative method for healthcare professionals to understand how infants perform in their home setting.

The AIMS home-video method is feasible for parents of typically developing children (Boonzaaijer et al., 2019). Parents appraised the AIMS home video method positively and considered home videos of added value in monitoring infants at risk in neonatal follow-up (Suir et al., 2022). The intra- and interrater reliability between the AIMS home videos made by parents and simultaneous observation on-site by a physiotherapist as a gold standard in 48 infants (1.5-19 months of age) was high, with intraclass correlation coefficients higher than 0.99, according to 94% of the parents, and 6% were excluded because of failure to follow the procedures. Videos were used to evaluate the inter- and intra-rater reliability among 3 trained testers, twelve physiotherapists familiar with the AIMS, who attended 2 training sessions of 3 hours led by experts (Boonzaaijer et al., 2017).

Home-based infant motor assessment provides opportunity and accessibility for families in resource-limited communities (Lima et al., 2022). The AIMS home-video method allows parents to record their child's motor behaviour at home, guided by instructions to upload the videos from their smartphone or camera through a computer via a web application specifically designed for this purpose (Boonzaaijer et al., 2019). The home-based infant motor assessment empowers parents to feasibly record their infant's motor behaviour at home that is not expensive for early developmental assessment. However, home-based infant motor assessment still has practical difficulties, such as requiring a special place and equipment for recording the home videos, parents' varied knowledge and interpretation of infant development, and the need for professional assessment and feedback on infant developmental status.

doi:10.6342/NTU202504540

1.4 Parental Perception of Infant Motor Assessment

Parents' perceptions of infant motor competence are the mental representations that fathers and mothers have of their infants compared to typically healthy developing infants (Hernández et al., 2011). Parents' perceptions of children's motor competence are important because the perceptions influence their interaction with the children and the results of developmental screening (Silva et al., 2017). Parents' perceptions may impact their interaction and engagement in children's motor activities, as their motivation to facilitate new motor abilities in acquiring their full potential. Thus, a well-developed personal parent competency is a fundamental mechanism promoting motor development engagement.

One of the most widely used parent reports, as seen by the size of its standardization sample, is the Ages and Stages Questionnaire. The ASQ is designed for assessing children aged 2 to 66 months among five domains: personal social, gross motor, fine motor, problem solving, and communication.

The Early Motor Questionnaire (EMQ) is a parent-report instrument specifically developed to capture infants' early motor abilities, including gross motor, fine motor, and perception–action skills, during the first two years of life. It was initially validated as a reliable and efficient tool for assessing motor development in infants aged 3 to 24 months, providing a cost-effective and unsupervised alternative to examiner-based assessments (Libertus & Landa, 2013). The EMQ shows good concurrent validity in comparison to the examiner-administered motor development measures the Mullen Scales of Early Learning (MSEL) ($r = 0.91 – 0.97$), but fair concurrent validity compared with the Peabody Developmental Motor Scales (PDMS-2) ($r = 0.40 – 0.47$) in children aged between 3 and 24 months (Libertus & Landa, 2013). These results

doi:10.6342/NTU202504540

suggest that parents can provide parental reports of their child's fine motor and gross motor development for early assessment.

Different from the EMQ, which is a parent-reported assessment of fine motor, gross motor, and perception-action development, the Maternal version of the Alberta Infant Motor Scale (MAIMS) is a parent-reported assessment that examines only the gross motor development (Abbott et al., 2000). The MAIMS was modified from the same 58 movements of the original version of AIMS for the purpose of maternal usability. Every item is presented with the pictures extracted from the original version, with simplified item instructions. Mother has to circle "Y" for Yes if they observed the movement or "N" for No if they never observed the movement (Piper & Darrah, 1994).

The MAIMS was examined for the validity of the mother assessment of two groups of infants, 30 full-term infants and 30 preterm infants, between 1 and 15 months of age. Comparison of the maternal perception of motor development with the physiotherapists' administration of the AIMS showed high sensitivity and specificity. The intraclass correlation coefficients between mothers' and physiotherapists' assessment results were 0.99 for the term group and 0.95 for the preterm group. Sensitivity, specificity, and positive and negative predictive values for the full-term mothers rating were all 1.00, but for the preterm mothers rating, sensitivity was 0.30, specificity 0.90, positive predictive value was 0.60, and negative predictive value 0.72 (Bartlett, 1993). The results of this research showed that mothers of full-term infants were accurate in detecting infants who had normal development and motor delay, while mothers of preterm infants were accurate in detecting those with normal motor development. These findings indicate the potential for utilizing parental perception reports as an infant motor assessment tool in clinical settings to identify infants who may require further

doi:10.6342/NTU202504540

assessment or intervention efficiently. However, their low validity results in preterm infants require further investigation and improvement.

Forty-three mother-infant dyads participated in the study of the relationship between the home environment and infant motor development. Using the MAIMS, mothers predicted their infants' motor development at eight months of age when the infants were five months of age. When infants were five months of age, mothers were asked to predict their infant's motor developments at eight months of age using the MAIMS. Then, infants were administered the HOME inventory and the AIMS by physiotherapists at eight months of age (Abbott et al., 2000). Although no statistically significant correlations between aspects of the parental expectation and motor infant development were found (Pearson's correlation, $r = 0.19$, $p = 0.23$), the forty-three mother-infant dyads mothers participated in this project had high expectations of eight-month motor performance and their infants had higher motor scores than normative samples on aspects of the home environment. These findings suggest that more supportive and stimulating home environments are associated with higher infant motor development scores (Abbott et al., 2000).

1.5 Artificial Intelligence for Infant Motor Assessment

Artificial intelligence (AI) provides a promising transformation of health care (Lanzagorta-Ortega et al., 2022). The AI research on adults using the Max Planck Institute for Informatics (MPII) Human Pose dataset consists of images taken from a wide range of real-world activities with full-body pose annotations (Wang et al., 2021). In recent years, as computer vision and AI algorithms have accelerated, there is increasing evidence that technology-assisted assessments of movement in infants' motor development are possible (Zhang et al., 2023). Healthcare professionals now have the opportunity to explore AI-based assessments of infants' motor development.

doi:10.6342/NTU202504540

AI technology, particularly convolutional neural networks (CNNs), has significantly advanced human pose estimation models. These models are now capable of accurately detecting limb joints, joint connections, and body positions, providing a detailed and precise understanding of human movement and posture (Moccia et al., 2020). AI models learn human movements from input data, recognizing joint positions in human skeletons (Chen et al., 2023). The skeleton model was used as an input vector for a shallow multilayer neural network (SMNN), which then performed movement classification (Reich et al., 2021). Human skeletons in videos are typically represented as sequences of joint coordinate lists, extracted by pose estimators (Duan et al., 2022). The use of skeleton-based models for human movement recognition has improved with the tracking of more joints, making movement analysis more accurate. Initially, 7- and 9-point models tracked basic joints such as the head, shoulders, elbows, wrists, hips, and knees, helping to assess basic human movements (Adde et al., 2021; Groos et al., 2022; Shrivastava & Pandey, 2017). As technology advanced, the skeleton model was added with 12 joints (hip center, spine, shoulder center, head, elbow left, wrist left, hand left, elbow right, wrist right, hand right, ankle left, and ankle right allowing for better tracking of human movements (Chawanpaiboon et al., 2019). Later, 17-point models included more joints, these skeleton-based action recognition models consist of 17 distinct joints or body parts: nose, eyes, ears, neck, right shoulder, right elbow, right wrist, left shoulder, left elbow, left wrist, right hip, right knee, right ankle, left hip, left knee, and left ankle, enabling the recognition of more complex human movements (Ojeleye, 2023). The most advanced 25-point models tracked even more details, such as facial features and fingers, allowing for better analysis of fine motor skills, including thumb tracking, end-of-hand tracking, and open and closed hand gestures (Reich et al., 2021; Shi et al., 2021). The increase in tracked points greatly improves the detection of human movement, as more data points lead to

doi:10.6342/NTU202504540

higher precision. The 17-point model is quite comprehensive for analyzing gross motor movements, while the 25-point model provides even more detailed information, making it highly effective for detecting fine motor skills.

Increasing evidence supports the potential of AI-assisted assessments of infant motor development (Zhang et al., 2023), providing healthcare professionals with opportunities to explore AI-based assessments. Several studies show that pose estimation and action recognition models have been successfully used in different populations. For example, a study on 12 3-month-old infants with and without complex congenital heart disease (CCHD) utilized MediaPipe's full-body model to detect arm and leg movements, overall, 499 leg and arm movements were identified, and the model had a PPV of 85% and a sensitivity of 94% with the Test of Infant Motor Performance and the Bayley Scale for Infant and Toddler Development (Rosales et al., 2024). Another study using video-based approaches for General Movement Assessment (GMA) has shown potential for detecting infant movements. One study utilized limb-pose estimation through a deep learning framework consisting of a detection and regression convolutional neural network (CNN). The study analyzed 3,200 frames of limb monitoring from preterm infants in a neonatal intensive care unit (NICU) and achieved strong results, with a high Dice Similarity Coefficient (DSC) of 0.94 (Moccia et al., 2020). DSC is a popular choice for comparing the agreement between the predicted segmentation against a ground-truth (Raina et al., 2023). Another study applied a CNN-based pipeline, validated on a dataset of 27,000 depth frames capturing preterm infants' spontaneous movements in the NICU. The results were compared with annotations provided by clinical or technical experts, serving as the ground truth. The pipeline achieved a median recall of 0.894 (Migliorelli et al., 2022). AI-based assessment is effective in classifying infant movements, but most studies focus on diagnosing specific

doi:10.6342/NTU202504540

conditions like cerebral palsy rather than analyzing the quality or variations in motor development, highlighting the need for further exploration of movement-based classification that examines detailed aspects of motor skills and development, along with age-based assessments to monitor developmental changes in motor skills according to age

1.6 AI with Mobile Application for Infant Motor Assessment

Recent advances in computer vision and deep learning have opened up new possibilities for extracting information from video recordings (Rezaei et al., 2019). Such technological innovations have led to an increase in attempts to refine recording techniques and to develop new tools adapted to different research and clinical needs. The application of smartphone-based solutions has been at the forefront because these devices allow for recordings in different settings, whether in the clinic or remotely, and help with accurate documentation of the timing of data acquisition (Marschik et al., 2023). Video-based assessments based on AI with machine learning hold promise to be widely accessible as they can be implemented on mobile devices (Rezaei et al., 2019). The experimental research demonstrated the possibility of expanding expert classification items from real-time images of younger infants captured by mobile devices using AI methods (Zhang et al., 2023).

The Movidea application, utilized in Italy in 2020, examined spontaneous movements through 300 videos from the NIDA database, involving 90 infants in total (Baccinelli et al., 2020). The Movidea software was developed using MATLAB version R2017a and its standard tools. It allowed for home-based analysis with single-camera setups and included automatic motion analysis, ensuring stable tracking and reliable outcome. The ICC coefficients reported were higher than 0.75 for all the features, indicating an excellent degree of agreement between the measures taken from the two operators.

doi:10.6342/NTU202504540

Smartphone technology, such as the In-Motion-App for home-based video recordings for automated tracking of spontaneous movements of high-risk infants, has an accurate method for early identification tested in a study across Norway, Denmark, and Belgium (Adde et al., 2021). The 86 parents returned 130 videos, performing two separate video recordings for their infant aged 12 to 13 weeks and 14 to 18 weeks post-term. All videos classified as 'GMA scorable' were consecutively assessed by one certified and experienced general movement observer who had passed advanced GMA courses under the General Movement Trust. The 7-point body tracker software correctly detected more than 80% of body key point positions. Most families found the instructions for filming their baby easy to follow, and more than 90% reported that using the instructions did not make them more worried about their child's development (Adde et al., 2021).

Video-based automatic motion analysis has been employed to identify infant motor development delays using the Pull to Sit (PTS) movement. The study aimed to develop an AI model using videos taken by mobile phone to assess infants' motor skills based on the PTS levels from the Hammersmith Infants Neurological Examination (HINE), and each examination was evaluated by a senior physiotherapist (Zhang et al., 2023) A total of 270 videos of 41 high-risk infants were taken by parents using a mobile device with the mean age while filming was a corrected age of 4.8 months old. An AI model with minimal environmental restrictions can provide a family-centred developmental delay screen and enable the remote monitoring of infants requiring intervention. By using videos captured on a mobile device with minimal limitations and AI methods, the accuracy of recognizing physiotherapist levels based on HINE is not inferior to previous automatic recognition systems under restricted filming environments. The average accuracies of the whole-body skeleton and key point models for level 0 were 77.7%

doi:10.6342/NTU202504540

and 88.1%, respectively (Zhang et al., 2023). The using of AI in mobile applications makes it easier for parents to assess and monitor their infants' motor skills remotely with good accuracy. However, infant motor development examined by pull-to-sit test movement alone could not predict motor development throughout an early age.

Although these applications offer valuable insights into infant motor development, they have certain limitations. For instance, GMA assesses only a limited range of movements in the supine position, and PTS focuses on a single movement and position that can only be observed in the early stages of development. As a result, they may not comprehensively capture motor development milestones across all positions throughout early childhood. Moreover, despite the use of automated analysis AI technology in the smartphone, the videos' results were still reviewed by trained assessors to ensure they met the required standards. This dependence on human evaluation, even with automation, shows that the process is not entirely automated and still requires human oversight for validation.

1.7 Preliminary Study of AI and Parental Perception for Infant Motor Assessment

Our previous study tried to develop and validate an AI model framework for movement recognition in full-term and preterm infants, as well as whether the pose estimation model and the skeleton-based action recognition model for adult movement classification are applicable and accurate for infant motor assessment. Our study prospectively evaluated 30 full-term infants and 54 preterm infants using the Alberta Infant Motor Scale (58 movements) and was recorded by 5 video cameras simultaneously in a standardized clinical setup. The movement videos were annotated for the starting time to the end, and the presence of movements by three physiotherapists. The result AI algorithm was accurate in classifying 31 movements in full-term

doi:10.6342/NTU202504540

and preterm infants from 4 to 18 months of age in a standardized clinical setup (accuracy = 0.91, recall = 0.91, precision = 0.91, and F1 score = 0.91) (Lin et al., 2024).

Another study was conducted to examine the concurrent validity of the AIMS-Short version (AIMS-S) with 31 movements in full-term and preterm infants when compared with the AIMS-Long version (AIMS-L) results with 58 movements in 44 full-term infants and 56 preterm infants during 4 to 18 months of age (Purwanto et al., 2023). The results demonstrated acceptable levels of concurrent validity compared with the AIMS with 58 movements results with agreement of 0.91, sensitivity of 0.93, specificity of 0.91, negative predictive value of 1.00, but low in positive predictive value 0.41. In 31 movements, "development not within the expected range" was defined when the infant showed absence of one or more essential items at the assessment age; in the AIMS assessment, "development not within the expected range" was defined when the infant's percentile was $<10^{th}$ centile % - 90% at assessment age.

Following up on the research on laboratory videos, seven more motions were added to the original 31 movements, for a total of 38 movements (Purwanto et al., 2023). The validation of machine learning has proceeded to 38 movements out of the 58 movements of the AIMS in full-term and preterm infants at age 4 to 18 months in a standardized clinical setup with an accuracy of 0.93, recall 0.92, precision 0.93, and F-score 0.92. Assessment of concurrent validity 38 movements with the original 58 AIMS movements across the 235 trials in our laboratory, and results showed high values in agreement 0.91, sensitivity 0.87, specificity 0.91, and negative predictive value 0.99, but low in positive predictive value 0.41. In 38 movements, "development not within the expected range" was defined when the infant showed absence of one or more essential items at the assessment age; in the AIMS assessment, "development not within the expected range" was defined when the infant's percentile was $<10^{th}$ centile at assessment age.

doi:10.6342/NTU202504540

We have further developed an application, *"Baby Go,"* to collect infant movement videos recorded by parents at home and examine parental perception of infant movements. Using AIMS 58 movements, 37 preterm infants and 54 full-term infants were prospectively examined for their movements at 4 to 18 months in this study. The result of the parental perception of movement videos recorded in full-term and preterm infants by the parents at four to 18 months of age under the home setup showed fair agreement with the physiotherapist's assessment results (60.3% (Hsiao, 2023). The parents uploaded 1.027 home videos, and the application of the AI algorithm for infant motor assessment on home videos showed an accuracy of 0.77, precision of 0.66, recall of 0.66, and F-score of 0.65 with the agreement of parental perception of infant movement with a physiotherapist's results from 59.8% to 61.3% (Hsiao, 2023).

There are a number of drawbacks to earlier research on the use of parental perception and AI in infant motor assessment. First, although data collection was conducted in the home environment by parents, the evaluation was accomplished by professionals. Second, the infant motor movements assessment in the application is still limited to a small number of movements and positions that do not reflect movements in infants' motor development throughout the early age. Third, the movement recognition model in our previous study of home videos still achieved a moderate level of accuracy compared to laboratory videos. Fourth, a discrepancy between professional and parent perceptions in another study of parental perception results points to a lack of parental education regarding normal infant motor development milestones and characteristics in the expected age range. Fifth, there is a need for an application that can be used as a screening tool in the early identification of infants' motor development.

Based on the gap above, this research was to develop an AI application based on home videos with more comprehensive infant movements and positions based on standardized

doi:10.6342/NTU202504540

assessment tools that can be used by parents with parental perception assessment of the infant's motor development, which included a parent education and provided developmental assessment and follow-up at the hospital.

1.8 Study Purposes

The purpose of this study was three-fold: (1) to upgrade the mobile application *"Baby Go"* to contain features of parental perception assessment, AI assessment of infant movements, and education full-term and preterm infants, (2) to examine the agreement between the parental perception and AI results from the application; and (3) to evaluate the agreement of the parental perception and the AI results with physiotherapists' labelling results.

1.9 Hypothesis

The corresponding hypotheses for these purposes were: (1) The upgraded "Baby Go" mobile application contains integrated features of parental perception assessment, AI assessment of infant movements, and education for full-term and preterm infants, (2) parental perception results agreed with AI assessment results from the "Baby Go" application, and (3) parental perception results and AI assessment results agreed with physiotherapists' labelling results of infant motor development.

doi:10.6342/NTU202504540

Chapter 2. Methods

2.1 Subjects

This study recruited full-term and preterm infants aged 3 to 18 months (corrected age for preterm infants) from the neonatal intensive care unit and the neonatal follow-up clinic of National Taiwan University Children's Hospital, Taipei, Taiwan. The inclusion criteria for full-term infants were gestational age 37-42 weeks and birth weight > 2,500 grams, and for preterm infants were gestational age < 37 weeks and birth body weight < 2,500 grams. The exclusion criteria for all infants were congenital and genetic abnormalities. All parents were informed of the study and signed a written consent form before participation. The hospital's human rights review committee approved the study (202012089INB and 202010031RINB)**.**

In our previous study, physiotherapists conducted the AIMS assessment on 84 infants while video recording their movements in the laboratory, resulting in 11,454 data samples. Using the *"Baby Go"* application, parents of 44 of the 84 infants who took part in the preliminary study video recorded their infants' movements bimonthly at home, contributing 1,895 data samples. In this study, machine learning will be based on 11,454 data samples from previously collected laboratory videos and 1,895 data samples from home videos. An additional 9,559 data samples will be collected through this study, requiring an estimated 94 infants to contribute home videos biweekly, accounting for an attrition rate of 20%.

2.2 Data Collection

We recruited a total of 91 infants, from October 2024 to June 2025 (with recruitment continuing until the end of June 2025), who were eligible for this study. Of these, 51 infants (32

doi:10.6342/NTU202504540

preterm and 19 full-term) had their parents sign the informed consent form and register the *'Baby Go'* application on their mobile phones. From this group, 45 parents of 29 preterm infants and 16 full-term infants uploaded at least one home video during the study period. Parents conducted parental assessments by regularly recording and uploading videos of their infants at home (Figure 1).

All data were collected using a mobile application called the *"Baby Go"* application, which was previously developed in an earlier project by a graduate student from the Department of Computer Science and Information Engineering at National Taiwan University. The application is available on the App Store and can be installed on both iOS and Android smartphones. After parents signed their consent form to participate in the study, a physiotherapist introduced them to the *"Baby Go"* application, including how to install, register, and use it. Parents then downloaded the application and registered a new account through the application administrator. To support their understanding, a written user guide was also provided.

The initial data collection involved gathering perinatal and demographic information through parental interviews, including the infants' sex, gestational age, date of birth, birth weight, maternal and paternal age, parental occupation, educational background, and primary caregiver. This information was entered by parents into the *"Baby Go"* application. The application also included a feature for assessing parental perception of infant movement, where parents observed and assessed the presence or absence of age-specific movements from 3 to 18 months in supine, prone, sitting, and standing positions.

Parents were instructed to video record their infants' movements at home every two weeks in the specified positions, following a standardized recording procedure to ensure good quality. These videos were uploaded via the *"Baby Go"* application from the age of 3 months

doi:10.6342/NTU202504540

until 18 months or until the infant began to walk. Graduate students from the Department of Computer Science and Information Engineering downloaded all data from the server, including infants' demographic information, parental perception assessments, and home videos, and anonymized the data to protect patient privacy and data security. Each week, data from the web portal was analysed and updated to include parental assessments based on the uploaded videos, automated movement classification results from the AI model in the application, and movement labeling conducted offline by physiotherapists as the gold standard.

2.3 Development of "Baby Go" Application

2.3.1 "Baby Go" Application Version 1.0

The first version of the *'Baby Go'* application was released on 2021-11-22 to upload home videos for purposes only, accompanied by an instruction manual for mobile phone video recording for parental use at home, with simple animations of 58 infant movements from the Alberta Infant Motor Scale (AIMS) (Figure 2).

In the first version, the application included a manual instruction feature to guide parents in recording home videos. Parents were asked to ensure that the videos clearly captured the infant's movements on a firm surface and a clean background. During home video recording, parents were instructed to maintain an optimal distance from the infant, hold the mobile phone steadily without shaking, record in a horizontal position while avoiding vertical angles, and adjust the video settings to ensure good video quality. Parents were also asked to ensure there was enough light so that the infant could be seen clearly in the video.

On 2022-03-29, *'Baby Go'* version 1.0 was revised into version 1.1, which contained features of infant movements accompanied by text description and picture demonstration to assist the parents in understanding the movements to be recorded and assessed. The parents

doi:10.6342/NTU202504540

could record home videos via the feature and then upload video files to mobile applications, or they could upload video files from mobile photo albums that they recorded with their mobile phone. All home videos from the parents were stored on the server of the Department of Computer Science & Information Engineering for data processing and analysis.

On 2022-10-22, the application was further revised into *"Baby Go"* version 1.2 to include more detailed instructions in the manual for mobile phone video recording and to add demonstration videos to make it more user-friendly. Furthermore, the descriptions of movement features for the animation pictures were revised to improve parents' understanding. *"Baby Go"* version 1.2 provided more detailed instructions for parents in video recording, which emphasized good quality of home video recording, such as recording in a horizontal view instead of a vertical view, ensuring stable camera movement, positioning the camera at an appropriate angle, and minimizing occlusion of the infant's body parts during recording.

Version 1.2 made revisions to the descriptions of some movements to help enhance parents' perception and interpretation of the movements to be selected and recorded. *"Baby Go"* version 1.2 also included a parental perception assessment checklist of 38 infant movements, designated for individual assessment aged from 3 to 18 months, with each age containing 2 to 5 movements. Parents were required to complete the checklist by ticking "Yes" if they observed their infant performing the movement and "No" if they did not, followed by recording and uploading videos.

"Baby Go" version 1.2 incorporated updated AI algorithms to classify 38 movements from home videos uploaded by parents. This AI-based application tracked and classified infant movements as present or absent using previously established algorithms, and determined whether the child's motor performance was "development within the expected range" or "development

doi:10.6342/NTU202504540

not within the expected range." Although AI integration was completed, this version was still under development and had not yet been used by parents. Version 1.2 also provided a parental education component to guide parents in conducting age-appropriate motor activities at home. We used *"Baby Go"* versions 1.0, 1.1, and 1.2 in our preliminary study primarily for video uploading purposes, but not for developmental assessment.

2.3.2 *"Baby Go"* Application Version 2.0

For this study, we revised *"Baby Go"* version 1.3 to version 2.0 to include features of parental perception reporting, AI-based movement assessment, developmental follow-up, and parental education. The parental education component was structured into four age groups: 4–6 months, 7–9 months, 10–12 months, and beyond 12 months, based on information established previously (鄭 & 楊, 2021).

On 2024-10-04, the application was revised to *"Baby Go"* version 2.0, which included more detailed instructions in the manual for mobile phone video recording and added demonstration videos to enhance user-friendliness. Additionally, the descriptions of movement features for the animation pictures were revised to improve parents' understanding (Figure 3).

In this study, version 2.0 of the *"Baby Go"* application was used for parents to record and upload videos of their infants from home, as well as to illustrate the AI classification results. However, in this version, the parental perception feature was not yet independent from the AI motor assessment and thus could not be completed directly within the application by parents. The parental perception assessment consisted of 38 age-appropriate movements for infants aged 3 to 18 months, including 12 prone, 5 supine, 8 sitting, and 14 standing movements. Each movement was presented with instructions and pictures to help parents easily recognize and identify them. This feature is already included in the current version, but since it is still combined

doi:10.6342/NTU202504540

with the recording and video uploading functions, parental perception data cannot yet be fully collected separately through the application. Therefore, starting from 2025-03-01, data on parental perception have been collected using a Google Form survey. This questionnaire included the full set of movements for infants aged 3 to 18 months, and parents were asked to indicate whether each movement was “observed” or “not observed.” The responses were then manually recorded by physiotherapists as parental perception data. These data were compared to physiotherapists’ labeling results based on videos uploaded by parents through the application, which served as the gold standard. Future versions of the application are planned to directly incorporate this parental perception checklist before parents begin recording and uploading videos, enabling streamlined parent-reported assessments and improving the efficiency of data collection.

“Baby Go” version 2.0 introduced several important new features to enhance user engagement and improve data collection quality. Additionally, reminder notifications were implemented to help parents maintain consistent video recording schedules, and a built-in video editor was added to allow users to trim and adjust their recordings before uploading, which improved the clarity and standardization of submitted videos.

Version 2.1, released on 2024-12-05, brought several improvements to the video uploading interface to streamline the user experience during video submission. These enhancements aimed to minimize confusion and resolve technical difficulties that parents had previously encountered. Clearer wording regarding the recommended recording duration was also provided, ensuring that parents understood the expected video length and could record accordingly for more accurate developmental assessments.

doi:10.6342/NTU202504540

On 2025-02-18, version 2.2 was introduced to include refinements of the assessment logic along with additional user support features. The logic was changed so that if a baby could perform a more advanced movement, any earlier movements that were not uploaded would automatically be considered passed. For example, if "Reaching from Extended Arm Support" was observed by the parent and classified as presence by the AI model, then "Forearm Supported (1)" and (2) would also be treated as presence. If "Sitting without Arm Support (2)" was present, earlier sitting movements, like "Sitting with Support" and "Sitting with Arm Support" would be considered present. If "Walks Alone" was present, earlier standing movements such as "Supported Standing" would be considered present. This change was made because some infants showed advanced development, and earlier movements did not show up anymore. In addition, the collection of parental perception data using the Google Form started in February 2025 and was implemented alongside the AI-based assessments.

The development of version 2.3 on 2025-03-08 was aimed at addressing technical challenges related to video orientation in the application and revisions in the interface and recording instructions. Specifically, the application was updated to accommodate videos recorded in vertical format, ensuring that they could be processed and evaluated correctly. This enhancement helped prevent errors in both AI analysis and the physiotherapist's labeling. To facilitate easier use for parents, additional revisions were made, including a simplified welcome interface, updated shooting guide instructions, lowered lens sensitivity, a parent self-evaluation feature, and monthly reminders. When the application indicated development not within the expected range, and this result was confirmed by the AIMS assessment conducted by physiotherapists, the infant was scheduled for developmental assessment and follow-up at the hospital.

doi:10.6342/NTU202504540

2.4 Data Collection

After completing the parental perception of infant movement performance, parents recorded their infant's movements according to the designated movements and then uploaded the video files via the *"Baby Go"* application (Figure 4). The AI algorithm was used to assess infant movements, whether the infant showed "development within the expected range" or "development not within the expected range."

If the AI results showed "development within the expected range," parents were advised to conduct the parental perception assessment and upload videos again in the following month. However, if the AI results showed "development not within the expected range," parents were advised to supplement the movement videos classified by the AI as absent within the time window of the assessment age (± 2 weeks).

Based on the results of the validation study conducted, it was found that there were more false positives than false negatives. To validate potential false-positive cases from the AI assessment, infants identified by the AI as having "development not within the expected range" were scheduled for an AIMS assessment by a physiotherapist at the laboratory during the follow-up period. To validate potential false-negative cases, infants identified by the AI as having "development within the expected range" were invited, if available, to undergo an AIMS assessment by a physiotherapist at the laboratory.

Parents continued uploading videos biweekly if the physical AIMS results indicated "development within the expected range." If the physical AIMS assessment showed "development not within the expected range," the infant was recommended for developmental assessment and follow-up at the hospital.

doi:10.6342/NTU202504540

To increase parents' motivation to continue participating in this study, small gifts were provided to each infant in the form of toys at the critical ages of 6 and 12 months, along with special rewards for accumulating the number of videos uploaded at milestones of 15, 30, and 38 videos. These gifts and special rewards were announced within the application.

2.5 The AIMS Assessment

The Alberta Infant Motor Scale (AIMS) is a standardized gross motor assessment for children from birth until 18 months of age that incorporates the neural maturation theory and the dynamical systems theory (Piper & Darrah, 2021). It contains 58 movements that assess the control and integrity of the antigravity muscles during observation of infant motor skills in prone (21 movements), supine (9 movements), sitting (12 movements), and standing positions (16 movements) (Piper & Darrah, 2021). Each movement is scored as one point if it is observed and zero point if it is not observed.

The normative values of the AIMS were established in 2,202 infants in Alberta, Canada, and were expressed as mean, standard deviation (SD), and percentile rank for the total score in every month of age, as well as time norms for the achievement of every movement (Piper & Darrah, 2021). Canadian infants, showing high reliability with a correlation coefficient of 0.99. The study also evaluated concurrent validity using 68 infants who were abnormal or at risk, by comparing the AIMS gross motor raw scores with both the Gross Motor Scale of the Bayley Scales of Infant Development and the Gross Motor Scale of the Peabody Developmental Motor Scale, yielding correlation coefficients of 0.93 and 0.95, respectively (Piper & Darrah, 2021). These results indicate that the AIMS demonstrates excellent reliability and strong concurrent validity with established motor assessment scales.

doi:10.6342/NTU202504540

The AIMS is used across the world, in particular to monitor the infant's gross motor development (van Iersel et al., 2020). The psychometric features of the AIMS have been studied in infants in several countries: Canada (intra-rater reliability r=0.99, inter-rater reliability r=0.99 and concurrent validity r=0.97) (Piper et al., 1992), Brazil (inter-rater reliability ICC=0.843 to 0.954 and internal consistency α = 0.90 (Valentini & Saccani, 2012), Greece (inter-rater reliability ICC = 0.99, 95% CI: 0.99–0.99) (Syrengelas et al., 2010), Japan (intra-rater reliability ICC = 0.86 - 0.99) (Uesugi et al., 2008), Taiwan (intra-rater reliability and inter-rater reliability: ICC = 0.97 – 0.99; the correlation of the AIMS scores with the BSID-II scores at 6 and 12 months: r=0.78 and 0.90, respectively; the prediction of the motor function at 12 months by the AIMS scores at 6 months: r=0.56 (Jeng et al., 2000)), and China (intra-rater reliability ICC = 0.81 - 0.99) and inter-rater reliability (0.98 - 0.99) (Wang et al., 2018).

In our previous AI study, 31 out of 58 movements of the AIMS were identified from machine learning with acceptable accuracy based on laboratory videos (Lin et al., 2024). Using the first arrangement of 31 movements, we conducted an assessment of concurrent validity of the 31 movements with the total AIMS assessment across the 235 trials in 44 full-term infants and 56 preterm infants during 4 to 18 months of age, demonstrating acceptable values of agreement 0.91, sensitivity 0.93, specificity 0.91, negative predictive value 1.00, and positive predictive value 0.41 (Purwanto et al., 2023).

Following up on the research on laboratory videos, seven more movements were added to the original 31 movements, for a total of 38 movements. Previously, some confusing movements sets were found during machine learning validation, such as rolling supine to prone with and without rotation and four points kneeling (1) vs. (2), therefore one of them

doi:10.6342/NTU202504540

with good accuracy was included in the 38 movements and subsequently the third arrangement was made by adding seven new movements and by moving several movements.

The seven new movements that were added included reciprocal crawling, reaching from extended arm support, supine lying (3), hands to feet, reaching with rotation in sitting, cruising with rotation and standing from quadruped movement and several movements that were moved included supported standing (2) moved from 6 to 7 months age group, sitting without arm support (2) at 11 months was also added to 12 months age group, standing from modified squat moved from 14 to 15 months age group, and the squatting moved from 15 to 16 months age group. The validation of machine learning against the labeling results of physiotherapists has proceeded to 38 out of the 58 movements of the AIMS in full-term and preterm infants at age 4 to 18 months in a standardized clinical setup with the accuracy of 0.93, recall 0.92, precision 0.93, and F-score 0.92 (Purwanto et al., 2025). Assessment of concurrent validity 38 movements with the original 58 AIMS movements across the 510 sessions in our laboratory showed agreement 0.99, sensitivity 0.92, specificity 0.99, negative predictive value 1.00, and positive predictive value 0.86 (Purwanto et al., 2025). The 38 movements consisted of 12 prone movements, 5 supine movements, 8 sitting movements, and 14 standing movements (Appendix 1). These 38 movements were then distributed to represent movements occurring within the age range of 3 to 18 months, with 2 to 5 movements allocated for each age group (Appendix 2). The final re-arrangement was subsequently incorporated into the parental assessment checklist in the "*Baby Go*" application.

2.6 Parental Perception Assessment of Infant Movement

The parental perception assessment of infant movement included 38 movements identified through machine learning in our previous study (Appendix 3). The parental perception

doi:10.6342/NTU202504540

assessment was designed to be conducted monthly from 3 to 18 months of age, following the scoring criteria of the original AIMS. Movements performed by 60–90% of normal infants were credited, following the standards of the Denver Developmental Screening Test – 2nd Edition (DDST-II) (Frankenburg et al., 1992) and Taipei Development Screening-II (Taipei-II) (Chen et al., 2020).

Every movement in the parental perception assessment checklist was presented with a picture extracted from the original version, along with simplified item instructions to facilitate easier identification by parents. The application provided demonstration videos of good quality, abstracted from laboratory recordings. Parents were asked to circle "Y" for Yes if they had observed the movement or "N" for No if they had never observed the movement.

Parents then completed the checklist in the parental assessment within the *"Baby Go"* application, ticking "Yes" if they had observed the movement and "No" if they had never observed it. In the parental perception assessment in *"Baby Go,"* "development within the expected range" was defined as the infant showing the presence of all movements appropriate for the assessment age, while "development not within the expected range" was defined as the absence of one or more movements within the normal age movement criteria. The assessment criteria were consistent with the physiotherapists' labeling results of the videos uploaded by parents, which served as the gold standard. It was necessary to carry out a preliminary study regarding parental perception of infant movement because, from using *"Baby Go"* in the previous version, parents reported that several movement descriptions were unclear, and they had difficulty determining exactly which movement to choose. Unclear movement descriptions were, such as in sitting without arm support (2) and cruising with rotation. It will be important to

doi:10.6342/NTU202504540

enhance the description revision of these movements for parents to better differentiate among them in the application.

We conducted a preliminary study in the laboratory using a parental perception assessment checklist on a total of 22 preterm and full-term infants aged 4 to 14 months to measure the agreement between parental perception assessment results and physiotherapists' assessments in the laboratory. The results showed a high agreement between parental perception assessment of infant movement and the physiotherapists' assessment on these movements (agreement = 0.82). In the age group of 6, 8, 12, 13, and 14 months, high agreement (1.00) was obtained for these movements for which may be frequently performed by infants at home. Low agreement (0.33) was found in sitting without arm support (2) and supported standing with rotation movement in the 10-month age group. This may be partly related to parents' misperception of the infant's leg movement in sitting without arm support (2). Therefore, we revised the text descriptions of some movements in the application features to reduce misperception and misinterpretation of the movements to be selected. Moderate agreement (0.60) was observed in supported standing (2) within the 4-month age group, which may have been due to this movement being rarely emphasized in childcare and difficult for a single parent to perform. Before the study started, these results guided how to better describe the movement features and explain them within the application to facilitate parental use. They also helped identify the ages of 4, 6, 8, 10, 12, 14, and 18 months as critical periods for parents to evaluate their infants' motor development.

Since the parental perception integration feature in the *"Baby Go"* application was still in the early stages of development, we also conducted a survey using a Google Form to collect data on parental perception. The form included a checklist tailored to the infant's age criteria,

doi:10.6342/NTU202504540

designed to help parents carry out developmental assessments at home more easily. Graduate students assisted parents by providing support whenever they encountered technical difficulties using the application, either through email or phone calls.

2.7 AI Assessment

2.7.1 AI Models

The mobile application *"Baby Go"* is available in the App Store and can be installed on iOS or Android mobile phones. A cell phone with HD 1920 × 1080 pixels and 30 fps was recommended for home recording to obtain high-quality images. The parents were asked to video record their infant's movements in prone, supine, sitting, and standing at home and then upload them via the application.

The data processing of video records consisted of three steps as below:

<u>Selection of Movement Video Recordings</u>

All video records were converted into mp4 format for storage and processing. Each selected video segment of movement was annotated by physiotherapists using the Action Interval Annotator V2 developed by our research team to annotate the start and end times and the presence of the movement. The physiotherapists were required to preview the video file with five camera views for the laboratory videos and one camera view for the home videos and then annotate the movements according to the AIMS scoring criteria and the labeling guidelines. Three physiotherapists with one to 20 years of pediatric experience received the labeling training course and achieved an agreement of > 0.85 with an experienced physiotherapist in video labeling (Hsiao, 2023; Lin et al., 2024).

doi:10.6342/NTU202504540

Establishment of the Pose Estimation Model

An action recognition model was established for action tracking and recognition through incremental machine learning on the collected infant movement video records. In the action recognition model, human skeletons in a video are mainly represented as a sequence of joint coordinate lists, where the coordinates are extracted by pose estimators (Duan et al., 2022), were considered. The skeleton-based action recognition models in our AI development (Yao et al., 2011), consists of 17 distinct body parts, including nose, eyes, ears, neck, right shoulder, right elbow, right wrist, left shoulder, left elbow, left wrist, right hip, right knee, right ankle, left hip, left knee, and left ankle (Ojeleye, 2023).

Training and Validation of the Action Recognition Model

The PoseConv3D model was used for movement recognition because of its robustness against potential noise from pose estimation and its good capability for spatiotemporal feature learning. It received an input of a sequence of 2D key points produced by the pose estimator on a video slice and converted them into a heat map using a Gaussian map (Duan et al., 2022). All the heat maps for each frame were stacked together sequentially to form a 3-dimensional heat map volume. Computation efficiency was further enhanced by cropping the heat map frames into the region of interest (which contained the subject) for the 2D heat maps and by performing uniform sampling for the third dimension (i.e., the temporal dimension) to reduce the sequence length. Given a list of target classes (movements), the PoseConv3D learned to predict the correct class from the key point sequence input.

In this study, the first AI model was trained solely on home video data with 1,396 data samples from home videos. However, this approach presented limitations, particularly because home video data tends to be more variable and limited in volume, making it harder for the model

doi:10.6342/NTU202504540

to achieve stable performance. To address this, a transfer learning strategy was adopted. In this approach, the second AI model pre-trained on the controlled laboratory dataset (474 videos with 7,234 data samples) was fine-tuned using the home video data (461 videos with 1,419 data samples). This allowed the model to leverage foundational knowledge learned from the lab environment and adapt it to the more unstructured conditions of home videos, which often include varied camera angles and motion artifacts such as shaking. This strategy not only improved accuracy but also increased training efficiency, which aligns with best practices in video-based action recognition (Carreira & Zisserman, 2017).

2.7.2 Data Collection for AI Assessment

Parents were asked to upload infant movement biweekly, if possible, through the *"Baby Go"* application for AI assessment. The accessible data included the application number, infant's age, assessment age, date and time (including start and end times of the videos), selected infant movement videos recorded and uploaded by parents, and the automatic identification results generated by the AI model.

Movements selected by parents, based on parental perception, were displayed in the application as movement numbers among the 38 movements listed in the age-based movement checklist. For automatic identification by the AI model in the application, the AI model would provide up to five alternative movements from the video uploaded by the parents, as more than one movement may be present in a single video.

2.8 Physiotherapists' Assessment

2.8.1 Video Labelling for AI Modeling

All the home video data were downloaded from the server by the graduate students of the Department of Computer Science & Information Engineering, National Taiwan University, and

doi:10.6342/NTU202504540

the files were sent to physiotherapists for movement labeling. Before commencing the home video labeling work, the three physiotherapists received training in labeling and were required to achieve an agreement greater than 0.85 with an experienced physiotherapist in video labeling (Hsiao, 2023).

All edited home videos were annotated by the physiotherapists using a labeling tool developed by the Department of Computer Science & Information Engineering team to mark the start and end times and the presence of the movements. Each movement was labeled as "perfect presence," "acceptable presence," or "partial presence." Perfect presence was defined as when the video fully displayed the movement characteristics, allowing the annotator to clearly recognize the item. Acceptable presence was defined as when the movement characteristics were not fully displayed due to inappropriate shooting angles or object occlusion, causing difficulty in differentiating between two similar items. Partial presence was defined as when the duration of a movement was shorter than the required time for the item. This definition was specifically applied to sitting movements, which normally required 15 seconds but could be labeled as partial presence if the movement was observed for at least three seconds.

Videos from the application, along with their labeled movements and video quality, were included in the data server, which was updated and evaluated weekly. Videos that could not be labeled either because they fell outside the 38 movement criteria, did not meet the labeling criteria, or were duplicates of previously uploaded videos were excluded from further analysis.

After the physiotherapists finished the home video labeling work, the labeled results were sent back to the graduate students of the Department of Computer Science & Information Engineering to serve as training datasets for further machine learning. All the labeled results served as the gold standard for model validation.

doi:10.6342/NTU202504540

2.8.2 Validation of AI Assessment Results

The labeling results by the physiotherapists were used as the gold standard for validation of the AI assessment results. A movement labeled as perfect, acceptable, or partial presence was considered presence and was used to compare with the AI classification results in this study. For reliability, two of the three physiotherapists performed the labeling independently. To ensure that physiotherapists' labeling was not influenced by AI results, the AI classification results were presented only as numeric codes without movement descriptions and were concealed during labeling. In case of any discrepancies between their assessments, the final decision was made through discussion by the annotators to reach consensus. In addition to labeling the movements, one physiotherapist also assessed the quality of the videos uploaded by parents, evaluating factors including video orientation, camera angle, camera stability, body part visibility, occlusion, and the presence of duplicate videos.

2.9 Movement- and Age-based Analysis

Infant motor assessment contained move- and age-based assessment in this study. Each assessment was examined for the agreement on movement selection in the three pairwise comparisons: the parents' vs. AI results, the physiotherapists' vs. AI results, and the parents' vs. physiotherapists' results.

2.9.1 Movement-based Assessment

The AI model accurately classified 38 movements in full-term and preterm infants aged 4–18 months in a clinical laboratory setting, and also showed validity when classifying the same 38 movements recorded by parents at home via the application in a previous study (Purwanto et al., 2025). Validation of the AI model for classifying 38 movements in full-term and preterm infants revealed an accuracy of 0.91, precision of 0.92, recall of 0.90, and F1 score of 0.91 with

doi:10.6342/NTU202504540

the laboratory videos and an accuracy of 0.84, precision of 0.84, recall of 0.77, and F1 score of 0.78 with the home videos. These movements were dispatched into age-based sets with two to five movements per age that showed high concurrent validity with the AIMS results (agreement = 0.99) (Purwanto et al., 2025).

The 38 movements comprised 12 prone position movements, 5 supine position movements, 8 sitting position movements, and 14 standing position movements (Appendix 1). These movements were grouped into two categories: static movements, referring to relatively stable postures maintained in a base of support for a certain duration (e.g., "Forearm support (1)", "Forearm support (2)"); and dynamic movements involving transitions between positions (e.g., "Prone to sitting", "Side-lying") or locomotor patterns involving displacement, such as "Reciprocal crawling (1)" and (2), and "Reciprocal creeping (1) and (2)".

The movement-based assessment was examined for the agreement between the parents' and the AI results, the agreement between the physiotherapists' and the AI results, and the agreement between the parents' and physiotherapists' results. Additionally, the movement-based assessment was also performed for the full-term and preterm infants, respectively, to examine whether differences existed in parental perception of infant movements.

2.9.2 Age-based Assessment and Its Concurrent Validity

The movements accurately classified by the AI model were organized into age-based sets monthly from 4 to 18 months (Appendix 2), according to the normative standards of the AIMS and the Denver Developmental Screening Test, 2nd Edition (Frankenburg et al., 1992), and the Taipei Developmental Screening-II (Taipei-II) (Chen et al., 2020). Before these age-based sets could be used for future infant motor assessment, their concurrent validity was examined using the AIMS results as the criterion measure.

doi:10.6342/NTU202504540

For the age-based assessment, delay was defined as the absence of one or more movements at the assessment age, while normal development referred to the presence of all movements appropriate for the assessment age in an individual child. For the physical AIMS assessment results, delay was defined when the child's percentile score fell below the 10th percentile, and normal development was determined when the child's score was equal to or above the 10th percentile (Piper & Darrah, 2021).

The data obtained from the application or collected through surveys were used to represent parental perception, based on parents' selection of "observed" or "not observed" for each movement when assessing their child at home. Physiotherapists conducted an age-based assessment by first reviewing the parental perception data, followed by offline video evaluations from the set of videos uploaded by parents within the corresponding assessment age group. The criteria for the age-based assessment were as follows: "development within the expected range" was defined as the infant demonstrating the presence of all age-appropriate movements, while "development not within the expected range" was defined as the absence of one or more movements based on the normative age-specific movement criteria.

2.10 Parent Education in Application

The *"Baby Go"* application version 2.3 included a parental education component to help parents understand how to perform age-appropriate gross motor activities at home. The gross motor activities and play guidelines were referenced from a book authored by the professor leading our infant movement laboratory, titled *Dancing with the Palm Fairy ~ Parenting Tips Before One Year* (鄭 & 楊, 2021).

Parental education was structured into four age groups: 4–6 months, 7–9 months, 10–12 months, and beyond 12 months. This educational guide was used by parents to facilitate gross

doi:10.6342/NTU202504540

and fine motor development activities while playing with their infants at home. For example, education for infants aged 4–6 months emphasized rolling exercises, pull-to-sit exercises, and supported sitting exercises, while encouraging parents to help the baby touch one hand with the other and move the baby's hands to touch knees or feet. At 7–9 months of age, the focus of gross motor activities was on sitting and sitting-to-standing exercises, while fine motor movements emphasized reaching for and holding safe objects, as well as practicing hand play through playing and singing with parents. For infants aged 10–12 months, gross motor skills focused on sitting-to-standing exercises and assisted walking practice, while fine motor training involved playing with digging toys and doodling with crayons on paper. For babies older than 12 months, parents focused on encouraging more independent walking in a safe environment and practicing opening book pages and writing in picture books or on whiteboards (鄭 & 楊, 2021).

2.11 Statistical Analysis

Full-term and preterm infants' perinatal and demographic characteristics were compared using t-tests for continuous variables and chi-square tests for categorical variables. For movement-based assessment, the agreement between the parental perception assessment, the physiotherapists' assessment results, and the AI results was examined by calculating the percentage of agreement out of the total number of assessments for individual movements. For age-based assessment, the agreement between the parental perception assessment, the physiotherapists' assessment results, and the AI results was examined by calculating the percentage of agreement for those trials with all movements observed out of the total number of trials at individual age.

doi:10.6342/NTU202504540

The concurrent validity of the age-based sets was examined in full-term and preterm infants at each age by comparing them against the physiotherapists' AIMS assessment results using agreement, sensitivity, specificity, positive predictive value, and negative predictive value (Carvajal & Rowe, 2010). Agreement was defined as the number of agreements divided by the total number of assessments. Sensitivity (SEN) referred to the percentage of infants classified as having motor delay by both the age-based set and the AIMS results among those identified with motor delay by the AIMS results. Specificity (SPE) was the percentage of infants classified as having normal development by both the age-based set and the AIMS results among those identified with normal development by the AIMS results. Positive predictive value (PPV) represented the percentage of infants classified as having motor delay by the age-based set who were also identified as delayed by the AIMS results. Negative predictive value (NPV) represented the percentage of infants classified as having normal development by the age-based set who were also identified as normal by the AIMS results. The criteria for acceptable validity were set at values greater than 0.8 for all indices (Šimundić, 2009).

doi:10.6342/NTU202504540

Chapter 3. Results

3.1 Infants' Perinatal Data

A total of 120 infants were projected for recruitment. By 2025-06-30, 91 infants were eligible for enrolment, and 45 parents (29 of preterm infants and 16 of full-term infants) signed the consent form and registered their infants in the *"Baby Go"* application for the study (Table 1-1). The proportion of female infants was similar between groups (full-term: 62.5%; preterm: 55.2%). Full-term infants had significantly higher birth body weight (2,949 ± 269.9 g vs. 1,756 ± 362.6 g) and gestational age (38.6 ± 0.8 weeks vs. 33.4 ± 2.1 weeks) compared to preterm infants (both $P < 0.05$). Apgar scores at 1 and 5 minutes were also significantly higher in full-term infants than in preterm infants (1-min: 8.2 ± 1.5 vs. 6.8 ± 2.2; 5-min: 8.9 ± 0.3 vs. 8.3 ± 1.3; both $P < 0.05$). Regarding respiratory distress syndrome (RDS), none full-term infants had RDS, whereas 55.2% of preterm infants presented varying severities: 55.2% in Grade 1, 13.8% in Grade 2, and 6.9% in Grade 3 ($P < 0.05$). Bronchopulmonary dysplasia (BPD) was observed in preterm infants only, with 69.0% classified as normal, 13.8% with mild, and 17.2% with moderate ($P < 0.05$).

The groups were comparable in the sex proportion, incidence of small for gestational age, periventricular leukomalacia (PVL), retinopathy of prematurity (ROP), necrotizing enterocolitis (NEC), and patent ductus arteriosus (PDA) requiring ligation. Neonatal hyperbilirubinemia was more common among preterm infants, with phototherapy required only in preterm infants.

3.2 Parents' Sociodemographic Data

doi:10.6342/NTU202504540

Of the 45 parents, paternal age was significantly lower in the full-term group (37.7 ± 3.6 years) compared to the preterm group (41.0 ± 7.2 years) ($P < 0.05$); whereas, maternal age was comparable between groups (Table 1-2). Full-time maternal caregiving was significantly more frequent in families of preterm infants (79.3%) than in those of full-term infants (50.0%) ($P < 0.05$). Family size was significantly larger among preterm families, averaging 4.9 ± 1.7 members versus 3.6 ± 1.0 members in the full-term group ($P < 0.05$). Likewise, the average number of children was higher in the preterm group (2.0 ± 0.7) than in the full-term group (1.4 ± 0.5) ($P < 0.05$). The groups were comparable in maternal age, parental education and occupation, annual household income, the proportion of infants primarily cared for by parents, and the proportion of firstborn infants.

3.3 Videos Uploaded for Movement- and Age-based Assessment

During the period from 2024-10-04 to 2025-06-30, 45 parents uploaded a total of 936 videos for movement-based assessment, comprising 534 movements (58%) from preterm infants and 393 movements (42%) from full-term infants (Table 2). The preterm group demonstrated a higher uploading rate at 5 to 10 months in movement-based assessment and at 5 to 9 months in age-based assessment. In the full-term group, a higher uploading rate was observed from 4 to 7 months in movement-based and from 5 to 10 months in age-based assessments. The higher upload rate for age-based assessments in the full-term group indicates better motor development.

For age-based assessment, preterm infants uploaded a total of 129 sets of movements, of which 24 (19%) were complete and passed, 17 (13%) were complete but did not pass, 55 (43%) were incomplete, and 33 (26%) were not uploaded at all (Table 3). In comparison, full-term infants uploaded 89 sets of movements, with 28 (31%) complete and passed, 12 (13%) complete but did not pass, 16 (18%) incomplete, and 33 sets (37%) not uploaded at all. Full-term infants

doi:10.6342/NTU202504540

had a greater proportion of complete sets that successfully passed the age-based assessment than preterm infants.

3.4 Movement-based Assessment Results

For the movement-based assessment, the agreement between the physiotherapists', parents' perception, and AI classifications results are presented for various versions (version 2.0, 2.1, 2.2, and 2.3), AI models (I and II), and the overall data are illustrated in Table 4. In general, the agreement between the physiotherapists' and parents' results was higher (75% to 82%), the agreement between the physiotherapists' and AI's classifications was moderate to high (65% to 85%), and the agreement between the physiotherapists' and AI's classifications was moderate (60% to 74%) across versions and AI models. The detailed agreement data for various versions and AI models are delineated in the following sections.

3.4.1 Movement-based Assessment Results for Application Version 2.0, 2.1, 2.2, and 2.3

A total of 78 videos uploaded by parents of 38 infants were assessed using the *"Baby Go"* application version 2.0. The agreement between the parents' and physiotherapists' results was 76%, the agreement between the physiotherapists' and AI classifications was 74%, and the agreement between the parents' and AI classifications was 68% (Table 5). High agreement (above 80%) was observed for several static movements in the supine position, including "Supine lying (3)" (90%), "Supine lying (4)" (83%), and 100% agreement for "Hands to knees" and "Hands to feet". Similarly, in the sitting position, "Sitting with arm support" also showed 100% agreement in parent-PT assessments. However, some movements showed low agreement (agreement below 60%), especially those that are similar and easily confused, such as "Forearm support (1)" and "Forearm support (2)" movement (50% parent-PT agreement respectively). Parents tended to be less accurate in identifying less mature with subtle different features,

doi:10.6342/NTU202504540

referring to those early movements. These results suggest that while AI classification aligns well with physiotherapy assessment, parental perception varies, especially for more static movements.

AI classification in in version 2.0, high agreement (above 80%) was observed between both PT and AI, and parent and AI, for several movements, particularly those that are static and clearly defined, such as supine-position movements including “Supine lying (3)”, “Supine lying (4)”, “Hands to knees”, and “Hands to feet” (ranging from 80% to 100%). In contrast, low agreement (below 60%) was found in several movements that typically involve parent’s assistance such as “Sitting with support” (PT–AI: 44%; parent–AI: 56%), “Pull to sit” (PT–AI: 0%; parent–AI: 0%), and “Reaching from extended arm support” (PT–AI: 0%; parent–AI: 0%). These findings highlight specific challenges faced by both AI models and parents in consistently identifying movements that involve external support or transitional postures, suggesting areas for further improvement in AI training and parent guidance.

A total of 199 videos uploaded by parents were assessed using the *“Baby Go”* application version 2.1. The agreement between parents’ and physiotherapists’ results was high at 82% (high agreement), the agreement between physiotherapists’ and AI classifications was 65% (moderate agreement), and the agreement between parents’ and AI classifications was 60% (moderate agreement) (Table 6). High agreement was observed for several static movements in the supine position, including “Supine lying (3)” (100% parent–PT agreement), “Supine lying (4)” (100%), “Hands to knees” (85%), and “Hands to feet” (86%). Similarly, static movements also showed improvement, “Forearm support (1)” (65%), and “Forearm support (2)” (86%). However, some movements showed low agreement, especially dynamic or transitional movements such as “Pivoting (50%), “Reciprocal crawling” (40%), and “Sitting without arm support (2)” (0%). These results suggest that while version 2.1 improved agreement for several

doi:10.6342/NTU202504540

previously challenging movements, such as “Forearm support (1) and (2)”, difficulties remained in identifying dynamic movements such as “Pivoting” and “Reciprocal crawling”.

AI classification in version 2.1, high agreement levels (above 80%) for certain well-defined supine movements remained consistent with version 2.0, including “Supine lying (3)”, “Supine lying (4)”, “Hands to knees”, and “Hands to feet”, with agreement ranging between 86–100% for both PT–AI and Parent–AI comparisons. Similarly, low agreement (below 60%) persisted for movements such as “Sitting with support” (26% for both PT–AI and Parent–AI) and “Reaching from extended arm support” (0% for both comparisons). With the addition of standing position movements in version 2.1, low agreement was observed for “Supported standing (2)” and “Supported standing with rotation”, both showing 0% agreement between PT–AI and Parent–AI. Further refinement of definitions and guidance is needed to enhance consistency across assessments.

A total of 39 videos uploaded by parents, containing 38 infant movements, were assessed using the *“Baby Go”* application version 2.2. The agreement between parents and physiotherapists was moderate (76%), the agreement between AI and physiotherapists was high (85%), and the agreement between parents and AI was moderate (74%) (Table 7). High agreement (above 80%) was observed for several static or postural movements, including “Supine lying (3)” (100% parent–PT agreement), “Supine lying (4)” (100%), “Hands to knees” (100%), “Hands to feet” (100%), and “Sitting with support” (100%). In addition, improvements were noted for previously inconsistent movements such as “Forearm support (1)” (67%) and “Forearm support (2)” (100%), as well as “Reaching from forearm support” (67%) and “Pivoting” (100%). Nevertheless, low agreement (below 60%) remained for some movements, especially those dynamic or transitional ones such as “Reciprocal crawling” (0%) and “Sitting

doi:10.6342/NTU202504540

without arm support (2)" (0%), as well as some standing movements, including "Supported standing (2)" (50%) and "Supported standing (3)" (0%). Although there were improvements in prone and supine movements, dynamic movements in the sitting position, and standing movements require further enhancement in future improvements.

In version 2.2, improvements in an agreement between PT–AI and Parent–AI assessments were observed in high agreement (above 80%), particularly in the supine position, where all relevant movements reached 100% agreement for both PT–AI and Parent–AI. Notably, agreement also increased for sitting position movements such as "Sitting without arm support (1)" and "(2)", each showing 100% agreement. Low agreement (below 60%) persisted for "Supported standing (2)", with both PT–AI and Parent–AI agreement remaining at 50%. However, the smaller sample size of 39 videos may limit the generalizability of these findings.

A total of 620 movements from videos uploaded by parents were assessed using the *"Baby Go"* application version 2.3. The agreement between parents and physiotherapists was moderate (75%), the agreement between AI and physiotherapists was moderate (65%), and the agreement between parents and AI was moderate (60%) (Table 8). Compared to version 2.2, *"Baby Go"* version 2.3 demonstrated improvements in the classification of standing movements. High agreement was observed in "Supported standing (2)" (100%), "Supported standing (3)" (82%), and "Stands alone" (100%), while "Pull to stand/stance" showed moderate agreement (73%). However, movements involving rotational components showed low agreement, such as "Reach with rotation in sitting" (52%) and "Cruising with rotation" (33%). These findings suggest that parents may experience difficulty identifying three-dimensional movements involving a rotation component, both in sitting and standing positions. While version 2.3 showed

doi:10.6342/NTU202504540

improvement in several static and postural movements, further refinement of movement definitions and enhanced parental guidance are needed to improve consistency in assessments.

In version 2.3, AI classification performance showed noticeable improvement across several motor domains. High agreement levels (above 80%) were particularly observed in the supine position, where all relevant movements achieved 80–100% agreement in both PT–AI and Parent–AI comparisons. In addition, moderate to high agreement was found in dynamic standing movements such as "Early stepping," "Walking alone," and "Squat", with agreement rates ranging from 75% to 100% for both PT–AI and Parent–AI comparisons. Moderate agreement (60% - 79%) was observed in prone and sitting positions for movements such as "Forearm support (1)" (PT–AI: 68%, Parent–AI: 72%), "Forearm support (2)" (PT–AI: 68%, Parent–AI: 72%), representing an improvement from earlier versions. However, low agreement (below 60%) persisted for movements that involved partial occlusion or rotational components. For example, "Reaching from forearm support" (PT–AI: 43%, Parent–AI: 38%) and "Reaching from extended arm support" (PT–AI: 19%, Parent–AI: 10%). Similarly, rotational movements such as "Reaching with rotation in sitting" (PT–AI: 45%, Parent–AI: 29%) and "Standing with rotation" (PT–AI: 43%, Parent–AI: 36%) also exhibited low agreement, suggesting that detecting rotational components remains a challenge for the AI model.

The whole data sample contained a total of 936 videos uploaded by parents assessed across *"Baby Go"* application versions 2.0 to 2.3. The overall agreement between the parents' and physiotherapists' results was moderate (77%), the agreement between AI and physiotherapists' results was 67% (moderate agreement), and 62% between parents and AI (moderate agreement) (Table 9). In the prone position, some movements remained confusing, with agreement ranging from moderate to low, such as "Forearm support (1)" (63%), "Extended

doi:10.6342/NTU202504540

arm support" (51%), and "Forearm support (2)" (40%), indicating persistent challenges in recognizing subtle prone variations. In contrast, static movements in the supine position consistently showed high agreement, reflecting improved clarity and ease of recognition by parents. In the standing position, the agreement was high in most of the movements, including "Stands alone," "Early stepping," "Standing from modified squat," "Standing from quadruped position," "Walks alone," and "Squat", demonstrating substantial improvements in recognizing upright postural and transitional tasks. However, movements involving rotational components remained difficult for parents to identify, with low agreement observed in "Reach with rotation in sitting" (53%) and "Cruising with rotation" (50%). These findings suggest that while parents increasingly recognized static and upright movements in prone and supine positions, further efforts are needed to enhance understanding of subtle prone variations and complex three-dimensional transitions involving rotational components.

From version 2.0 to 2.3, high agreement levels were observed in both supine and dynamic standing positions. In the supine position, movements such as "Supine lying 3," "Supine lying 4," "Hand to knee," and "Hand to feet" achieved 80–100% agreement in both PT–AI and Parent–AI comparisons. Moderate agreement was found in prone positions, particularly for "Forearm support (1)" (PT–AI: 68%, Parent–AI: 78%) and "Reciprocal creeping (1)" and "(2)" (PT–AI: 76%, Parent–AI: 72%). In contrast, low agreement (below 60%) persisted for movements such as "Reaching from forearm support" (PT–AI: 52%, Parent–AI: 45%) and "Reaching from extended arm support" (PT–AI: 17%, Parent–AI: 8%), likely due to the movements being partially occluded by the parent during video recording. Additionally, "Supported standing (2)" showed low to moderate agreement with PT–AI agreement at 23% and Parent–AI agreement at 80%. These findings indicate that while AI classification accuracy has

doi:10.6342/NTU202504540

improved over time, particularly in supine, prone, and dynamic standing movements, challenges remain in detecting occluded, rotational, or visually complex movements.

3.4.2 Movement-based Assessment Results on AI Model I

During the initial phase, from 2024-10-25 to 2025-01-13, a total of 194 data samples were assessed using the AI model I, which has been trained on 1,396 data samples from home uploaded through the first version of the *"Baby Go*" application (Table 10). The agreement between parents and AI was 64%, the agreement between AI and physiotherapists was 70%, and the agreement between parents and physiotherapists was 80%. High agreement (above 80%) was observed for several prone movements, including "Four points kneeling (1)", "Reaching from extended arm support", and "Four points kneeling to sitting", all achieving 100% agreement. Similarly, "Sitting with support", "Pull to sit", and "Sitting with arm support" also demonstrated high agreement, ranging from 96% to 100%. In contrast, dynamic movements such as "Reciprocal crawling" and "Sitting without arm support (2)" showed low agreement, with only 33% and 50%, respectively, while "Controlled lowering through standing" also had low agreement at 50%, reflecting challenges in consistent identification. Standing position movements involving support, such as "Supported standing (2)", "Supported standing (3)", and "Pull to stand with support" exhibited high agreement at 100%, reflecting reliable recognition by parents. However, more dynamic standing activities were either not recorded or lacked sufficient data.

AI classification using Model I, high agreement (above 80%) was observed in supine position, such as "Supine lying 3," "Supine lying 4," "Hand to knee," and "Hand to feet" for both PT–AI and Parent–AI comparisons. In the Prone position, moderate to high agreement was found for some prone movements such as "Forearm support (1)" and "(2)", with PT–AI

doi:10.6342/NTU202504540

agreement ranging from 65% to 90%, and Parent–AI agreement from 68% to 100%. Notably, "Reciprocal creeping I and II" achieved perfect agreement (100%) in both PT–AI and Parent–AI comparisons. In contrast, dynamic movements in the sitting position such as "Sitting without support I and II" showed low agreement for both PT–AI and Parent–AI agreement at only 50%. However, standing movements, data collection was still incomplete, and therefore, no analysis could be conducted. Overall, Model I performed better in identifying movements that were symmetrical, static, or clearly visible, particularly in prone and supine positions. In contrast, the model showed challenges in classifying movements that involved occlusion, rotation, or dynamic postural transitions, especially in the sitting and standing positions.

Following 2025-01-14, the *"Baby Go"* application implemented an updated AI model II, trained on a combination of 474 videos with 7,234 data samples collected from the laboratory, and 461 videos with 1,419 data samples from home, collected through the first version of the *"Baby Go"* application. This retraining aimed to improve AI performance in recognizing infant movements under the more variable conditions typical of home recordings. A total of 742 movements across 38 movement sets were analyzed (Table 11). High agreement was observed for most supine movements, including "Supine lying (3)" (80%), "Supine lying (4)" (96%), "Hands to knees" (95%), and "Hands to feet" (96%), indicating high consistency between parents and physiotherapists in assessing these static supine position. Similarly, sitting-related movements such as "Sitting with support", "Pull to sit"), and "Sitting with arm support" also demonstrated high agreement, reflecting reliable parental recognition of these sitting movements. In the standing position, moderate to high agreement was found in movements like "Supported standing (2)" (95%), "Supported standing (3)" (75%), "Pull to stand with support" (71%), and "Pull to stand/stance" (73%). However, low agreement was noted in more dynamic movements,

doi:10.6342/NTU202504540

particularly “Sitting without arm support (2)” (38%) and “Cruising with rotation” (50%), indicating ongoing difficulties for parents in accurately identifying dynamic movements involving greater independence or rotational components. These results suggest that although the updated AI model improved adaptability to home video variability, overall accuracy did not significantly increase.

AI classification using Model II demonstrated varied levels of agreement across different positions. High agreement was observed in several symmetrical and static movements. In the supine position, movements such as “Hands to feet” achieved perfect agreement (PT–AI and Parent–AI: 100%), and “Supine lying (3)” reached 100% for PT–AI and 80% for Parent–AI. In the prone position, “Reciprocal creeping (1)” and “(2)” showed moderate agreement for PT–AI (61%) and Parent – AI agreement (71%). In the sitting position, movements like “Sitting with propped arms” and “Sitting with arm support” reached moderate to high agreement, with PT–AI between 78–79% and Parent–AI between 64–78%. Similarly, in standing positions, static movements such as “Supported standing (2)” and “Stands alone” demonstrated high agreement, ranging from 82% to 100% for PT–AI. In contrast, low agreement was seen in movements that involved occlusion, rotation, or dynamic changes. For example, “Reaching from extended arm support” was difficult to classify, with PT–AI agreement at 19% and Parent–AI at 10%, likely because the movement was blocked during recording. Similarly, complex sitting movements like “Reaching with rotation in sitting” and “Sitting to four-point kneeling” had low agreement—PT–AI between 44% and 46%, and Parent–AI between 28% and 58%. In standing, dynamic and rotational movements such as “Pull to stand/stance,” “Cruising with rotation,” and “Supported standing with rotation” also showed low agreement, with PT–AI from 17% to 50% and Parent–AI from 25% to 36%.

doi:10.6342/NTU202504540

These results suggest that although the updated AI model improved adaptability to home video variability, overall accuracy did not significantly increase. Challenges such as incorrect video orientations, dynamic and complex infant movements, and inconsistent recording conditions likely continued to affect AI classification performance.

3.4.3 Technical Problems in Data Analysis

Table 12 presents the agreement between parents, the AI system, and physiotherapists based on 303 videos recorded in vertical (portrait) orientation. This represents a frequent technical issue where parents recorded videos in portrait mode rather than the recommended horizontal (landscape) orientation. High agreement was observed for several movements, even when videos were recorded in portrait mode. For example, “Supine lying (3)” (100%), “Hands to knees” (83%), and “Hands to feet” (100%) all demonstrated high parent–PT agreement. Sitting-related movements also showed high agreement, “Sitting with propped arms” (100%), and “Pull to sit” (80%). Notably, “Sitting without arm support (2)” reached moderate agreement (75%) despite the suboptimal video framing. However, low agreement (below 60%) was noted in more dynamic movements, including “Reaching from extended arm support” (11%), “Reciprocal crawling” (33%), and “Sitting to four points kneeling” (23%). The total agreement between parents and physiotherapists was 77%, indicating relatively consistent human assessments despite the portrait recording format.

The AI model I showed particular difficulty detecting dynamic movements involving changes in position and rotation. For example, “Four points kneeling to sitting” and “Reciprocal creeping (1)” both demonstrated only 38% agreement between AI and human raters, reflecting challenges in recognizing transitional motor patterns. Similarly, rotational movements such as “Cruising with rotation” had 50% agreement, while “Standing from modified squat” showed 0%

doi:10.6342/NTU202504540

agreement, highlighting the AI's struggle to interpret movements requiring complex body alignment and rotation in portrait videos. These findings underscore a critical technical limitation: portrait-mode adversely affected the AI's ability to detect movement direction and body alignment accurately. While human raters can mentally adjust for video orientation, the AI model currently has difficulty adapting to changes in camera angle and orientation in three-dimensional movements. This analysis included 142 videos identified as recorded with incorrect infant orientation (i.e., vertically flipped or rotated incorrectly) in the application (Table 13). High agreement was observed for several static and transitional movements despite incorrect video orientation, including "Supine lying (3 and 4)", "Hands to knees", "Hands to feet", "Pivoting", "Four points kneeling (1)", "Four points kneeling to sitting", "Pull to sit", and "Reaching from extended arm support." Moderate agreement was found for movements such as "Extended arm support" (71%). However, some movements showed low agreement, particularly dynamic actions like "Reciprocal crawling" (40%), "Reciprocal creeping (2)" (50%), "Sitting without arm support (2)" (33%), "Reach with rotation in sitting" (33%), and "Cruising with rotation" (0%). Overall, the total agreement between parents and physiotherapists was 80% (high agreement), indicating that human raters were able to consistently recognize infant movements despite orientation errors.

The AI model II demonstrated particular difficulty with several movements under these conditions. For instance, "Reaching from extended arm support" showed 0% agreement with physiotherapists, and "Pull to stand with support" showed 0% agreement with parents. Dynamic postural transitions and rotations, such as "Cruising with rotation" and "Pull to stand/stance," also showed 0% AI agreement, further highlighting the AI's challenges in analyzing dynamic movement patterns when videos were misoriented. In contrast, static supine movements such as

doi:10.6342/NTU202504540

“Supine lying (3)” (75–100% AI agreement), “Supine lying (4)” (63–70%), “Hands to knees” (100%), and “Hands to feet” (100%), as well as static prone positions including “Pivoting” (100%) and “Four points kneeling (1)” (100%), generally maintained higher AI agreement with both physiotherapists and parents. These static movements appeared more stable in classification, whereas dynamic movements involving changes in posture consistently showed substantially lower agreement. Overall, these findings underscore the necessity of incorporating automatic detection and correction of video orientation within the application prior to AI analysis. Encouraging parents to record videos with correct infant orientation and implementing orientation correction features will be essential to improve AI performance and reliability in future assessments.

The analysis of 438 videos that were recorded in portrait mode and/or incorrect orientation showed agreement rates of 74% between physiotherapists and parents, 52% between physiotherapists and AI, and 48% between parents and AI. These findings indicate that while videos with portrait mode and/or incorrect orientation resulted in moderate agreement between parents and physiotherapists that may have contributed to lower agreement between physiotherapists and AI, as well as between parents and AI**.**

This analysis included 498 videos recorded in proper horizontal orientation, excluding those with portrait mode or incorrect infant orientation (Table 15). High agreement was observed for several position in sitting and supine movements, including “Sitting with support” (97%), “Pull to sit” (95%), “Sitting with arm support” (95%), and “Sitting without arm support (1)” (100%), all demonstrating high parent and physiotherapist agreement. Similarly, static postural movements in the supine position, such as “Supine lying (4)” (86%), “Hands to knees” (94%), and “Hands to feet” (95%), showed high agreement ranging from 86% to 95%. Additionally,

doi:10.6342/NTU202504540

standing movements such as “Stands alone,” “Early stepping,” “Standing from modified squat,” “Standing from quadruped position,” “Walks alone,” and “Squat” all exhibited consistently high agreement at 100%, reflecting reliable recognition of standing position. In contrast, movements involving rotation during standing, such as “Supported standing with rotation” (50%) and “Cruising with rotation” (0%), displayed low agreement, indicating ongoing challenges in accurately identifying dynamic rotational movements. Overall, the total agreement between parents and physiotherapists increased to 79%, reflecting moderate to high agreement and indicating improved consistency in recognizing infant movements after excluding videos with portrait mode and/or incorrect orientation.

When all videos from application version 2.0 to 2.3 were included, agreement rates were 74% for physiotherapist–parent, 52% for physiotherapist–AI, and 48% for parent–AI, indicating moderate parent–PT agreement but lower AI performance. After excluding videos with portrait mode or incorrect orientation, overall agreement improved for all comparisons: physiotherapist–parent increased to 79%, physiotherapist–AI to 79%, and parent–AI to 73%. This demonstrates that removing portrait mode or incorrectly oriented videos significantly enhanced AI classification accuracy. To improve future assessment accuracy, parents should be reminded to record videos in horizontal orientation for optimal AI analysis. Additionally, the application could include automated features to detect and correct vertical recordings or further develop the AI model to reliably recognize infant movements even in portrait-mode videos.

3.4.4 Parental Perception for Movement-based Assessment

Parental perception data were analyzed based on movement-based assessments conducted using the *“Baby Go”* application. A total of 936 assessments were included, comprising 393 observations from full-term infants and 543 from preterm infants, covering 38 different infant

doi:10.6342/NTU202504540

movements. Table 16 presents the results, including all video data uploaded by parents, regardless of orientation. Overall, 78% of full-term assessments and 76% of preterm assessments were reported as “observed,” indicating a generally high level of parental recognition across groups. Certain movements, such as “Forearm support (1)” and “Reciprocal creeping (1),” had higher rates of “not observed” responses, possibly due to their similarity to other movements or the infants having progressed to more advanced motor skills. In contrast, movements like “Reaching from forearm support,” “Pivoting,” and “Sitting with support” were consistently reported as observed.

Table 17 presents the analysis after excluding videos recorded in portrait mode or with incorrect orientation, reducing the total assessments to 498 (284 from full-term and 214 from preterm infants). After these exclusions, the proportion of “observed” movements remained similar: 78% in the full-term group and 79% in the preterm group. Parental perception appeared consistent across different infant movements, with no clear trend indicating higher or lower recognition in either group. These findings support the feasibility and reliability of using parent-reported assessments via a mobile application to monitor infant motor development in the home environment.

Table 18 presents parental perception data in preterm and full-term infants, collected via Google Form between 2025-03-01 to 2025-06-30. A total of 202 movement observations were recorded, consisting of 131 observations from preterm infants and 71 from full-term infants. In the preterm group, 123 out of 131 observations (94%) were reported as observed, while 8 observations (6%) were measured as not observed. Most movements were observed by the parents, including "Forearm support (2)" (100%), "Pivoting" (100%), "Four-point kneeling (1)" (100%), and "Reciprocal crawling" (100%). Some movements showed slightly lower observation

doi:10.6342/NTU202504540

rates, such as "Forearm support (1)" (83% observed), "Hands to knees" (67%), "Hands to feet" (86%), and "Early stepping" (67%). Notably, "Forearm support (1)" had a 17% rate of non-observation, indicating some variability in parental perception of this movement. In the full-term group, 68 out of 71 observations (96%) were reported as observed, with only 3 observations (4%) as not observed. High observation rates were recorded for most movements, including "Forearm support (2)", "Rolling supine to prone with rotation", "Supported standing (2)", and "Cruising", all at 100% observed. A lower rate of observed movements were noted in "Extended arm support" and "Pull to stand with support" (50%).

Overall, combining both groups, 191 out of 202 movement assessments (94.5%) were perceived as observed, while 11 assessments (5.5%) were reported as not observed. These results suggest that the majority of expected movements in both preterm and full-term infants were successfully perceived by parents using the structured movement checklist provided in the Google Form survey.

3.5 Age-based Assessment Results

Table 19 includes 80 trials with complete movement sets from infants aged 3 to 18 months, including those uploaded in portrait mode or with incorrect orientation. The parental perception and AI each classified the individual infant's development as either “within the expected range” or “not within the expected range” according to age. The parents', AI, and physiotherapists' results were compared pairwise using the percentage of agreement. The agreement between the parents' and AI results was 64%, while the agreement between the AI and physiotherapists' results was 78%, and the agreement between the parents' and physiotherapists' results was 95%. These results suggest that parents, when guided by the application, could assess infant movements with a high level of agreement with physiotherapists.

doi:10.6342/NTU202504540

Performance varied across age groups: high agreement among all three comparisons was observed in infants aged 9 to 12 months., but low agreement was found in younger infants, such as at 3 months, where parent–AI agreement was only 43%. Similar patterns were noted at 8 months, indicating that age may influence assessment consistency.

Table 20 shows results after excluding videos recorded in portrait mode or with incorrect orientation, reducing the dataset to 61 trials. The agreement improved: parent–AI agreement increased to 74%, AI–physiotherapist agreement rose to 82%, and parent–physiotherapist agreement reached 97%. The improvement was especially pronounced in younger infants, such as at 3 months, where parent–AI agreement reached 100%. These findings highlight that correct video orientation is essential for reliable AI analysis and for improving the consistency of assessments across observers and age groups.

3.5.1 Parental Perception for Age-based Assessment in Full-term and Preterm Infants

Table 21 presents parental perceptions of age-based assessments for full-term and preterm infants based on data collected using the *"Baby Go"* application. A total of 81 trials were included (41 from preterm infants and 40 from full-term infants), covering infants aged 3 to 18 months. Parents were asked to indicate whether specific age-appropriate movements were fully observed or only partially observed. Overall, parental perception was high, with 95% of movements reported as fully observed in both full-term and preterm groups. Observation rates remained consistently high across most age groups, with full observation reported at 100% for several months, including 3, 4, 5, 8, 9, and 10 months in both groups. Minor reductions in full observation appeared at 6 months (83% full-term, 100% preterm) and 7 months (86% full-term, 100% preterm).

doi:10.6342/NTU202504540

To address video quality concerns, assessments excluding videos uploaded in portrait mode or with incorrect orientation were analyzed separately (Table 22). After these exclusions, 61 assessments remained (36 full-term and 25 preterm). The proportions of fully observed movements were similar to the full dataset, at 94% for full-term and 96% for preterm infants. These findings suggest that parental perception of infant motor development remained consistent regardless of video orientation issues.

Parental perceptions of infant motor development were also gathered through an age-based assessment using Google Forms, conducted between March 9 and June 30, 2025. There were 51 responses, comprising 17 trials from parents of full-term infants and 34 trials from parents of preterm infants. One duplicate response was removed, resulting in 50 valid data entries (Table 23). Overall, 82% of full-term parents and 79% of preterm parents reported that their infants had achieved all age-appropriate movements. The remaining 18% of full-term and 21% of preterm respondents reported only partial movement observation, which may suggest slower development or incomplete milestones in a minority of cases.

Overall, parental perception of age-based movement assessments was consistently high across all data sources and age groups, with only minor variations. There were no clear differences between full-term and preterm groups in their ability to recognize expected motor development milestones.

3.5.2 Concurrent Validity of Parental Perception for Age-based Assessment Compared with PT Labelling Results

A validity analysis was conducted to evaluate the agreement between parental perception and physiotherapists' labeling results for age-based assessment as either "within the expected range," indicating typical development, or "not within the expected range," suggesting possible

doi:10.6342/NTU202504540

motor delay. Using a dataset of 61 infants aged 3 to 18 months, key validity indices were calculated based on the comparison of both assessments (Table 24). The overall agreement rate between parental perception and physiotherapists' labeling results was 0.93, sensitivity of 0.86, specificity of 0.96, PPV was 0.86, and NPV was 0.96, suggesting high degrees of agreement between parental perception and physiotherapists' labeling results.

3.6 Concurrent Validity of AI Model for Age-based Assessment Compared with AIMS Assessment Results

To evaluate the concurrent validity of the AI-parent assessment model, a subset of 14 infants underwent laboratory-based motor assessments using the AIMS. The AIMS provides percentile scores indicating motor development relative to age norms, with infants scoring below the 10th percentile categorized as "below the expected range," and those at or above the 10th percentile as "within the expected range" (Table 25). These results were compared with combined parent-AI assessments conducted at home using the *"Baby Go"* application. Of the 14 infants assessed, eight (57.1%) were identified as normally developing by both methods, classified as true negatives. These included infants ranging from 3 to 16 months old, both preterm and full-term. Two infants (14.3%) were identified as having motor delays by both assessments, classified as true positives, with AIMS scores below the 10th percentile confirming the concerns flagged by the AI-parent model.

The analysis of the validity of AI results with the onsite AIMS assessment showed an overall agreement of 0.79 compared with the AIMS assessment results (Table 26). Sensitivity and NPV reached 1.00, meaning all infants with motor delays were correctly identified and no cases were missed. However, the specificity of 0.75 and PPV 0.40 indicate that some typically developing infants were misclassified as at risk of developmental delay.

doi:10.6342/NTU202504540

Chapter 4. Discussion

4.1 Development and Improvement of "Baby Go" Application

The *"Baby Go"* application has progressed substantially from its initial version to the second version, with enhanced feasibility and usability by parents of preterm and full-term infants. The majority of parents (90%) downloaded and registered the application, uploaded at least one video, with engagement rates of 93.5% among parents of preterm infants and 84.2% among parents of full-term infants. Between 2024-10-04 and 2025-06-30, the application collected a total of 936 videos for movement-based assessment and 80 sets of movements for age-based assessment from preterm and full-term infants aged 3 to 18 months. The findings indicate an improvement in parental engagement in the updated versions, which reflects the added value of integrating AI-based analysis, parental assessments, and educational support to foster higher participation in developmental monitoring.

When compared to the *"Baby Go"* version 1, the improvements in both engagement and functionality are evident. The initial prototype of the *"Baby Go"* application version 1.0 was developed primarily as a video-uploading tool for collecting infant movement data in the home environment for subsequent machine learning and the development of an AI model in preparation for remote infant movement classification (Hsiao, 2023). During the first feasibility study, 122 parents registered for the application, and 67.2% (n = 72) uploaded at least one video, with a lower engagement in parents of preterm infants (53.2%) than in parents of full-term infants (76.0%) (Tsai et al., 2025). Compared to other AI-based mobile applications for infant motor assessment, the *"Baby Go"* application version 2 showed higher parental use than for the existing video-based application such as the *NeuroMotion* (64.5%) (Svensson et al., 2021), the

doi:10.6342/NTU202504540

Baby Moves (70%) (Kwong et al., 2019), and the *In-Motion* (80.2%) (Adde et al., 2021). These findings suggest that combining AI analysis, parental perception, developmental follow-up, and educational support in version 2 contributed to higher parental participation. Previous research has shown that parents report higher participation when using applications that provide structured developmental assessments and clear instructions (Spittle et al., 2016), as well as interpretive feedback and meaningful developmental information, rather than functioning solely as data collection tools (DeWitt et al., 2022).

Parental engagement was higher among parents of preterm infants (93.5 %) compared to those of full-term infants (84.2 %). This difference likely reflects heightened parental concern and attentiveness toward preterm infants, who face greater risks of developmental challenges. Parents of preterm infants may be more concerned and motivated than those of full-term infants to engage with mobile applications due to their heightened need for reliable developmental guidance and emotional reassurance, as previously been reported by improved maternal confidence and perceived support when using a tailored parenting application for premature infants (Ahn et al., 2023).

While prior studies primarily focused on movement-based assessments to classify whether specific observed motor behaviors were typical or atypical, the present study advanced this approach by incorporating an age-based assessment of infant motor development. Specifically, the *"Baby Go"* application not only classified whether individual movements were typical (movement-based assessment), but also evaluated whether the infant was within the expected developmental age range, with all uploaded movement videos classified and confirmed by the AI (age-based assessment). The application included structured recording instructions with visual cues and age-specific movement groupings, enabling parents to assess and document

doi:10.6342/NTU202504540

their infant's development every month from 3 to 18 months of age. Both movement- and age-based features may enhance parental engagement in using this video-based application for infant motor assessment under a home environment (Migliorelli et al., 2022).

4.2 Movement-based Assessment Results across Versions and AI Models

The *"Baby Go"* application demonstrated a slight improvement in the assessment agreement across versions. Version 2.0 achieved 76% parent–physiotherapist agreement, which increased to 82% in version 2.1 following interface and instruction enhancements. Version 2.2 maintained a 76% agreement after algoritmic updates and added guidelines, while version 2.3 sustained a 75% agreement by resolving video orientation issues. These results indicate moderate to high validity of the application for home-based infant motor development monitoring. The findings suggest that iterative improvements in the *"Baby Go"* application have enhanced user satisfaction and assessment accuracy, aligning with similar studies that emphasize the importance of parental involvement in monitoring infant development through technology (Boonzaaijer et al., 2019). This reinforces the notion that user-friendly applications can facilitate better communication between parents and healthcare professionals, ultimately supporting the infant's development in a home setting (Adde et al., 2021).

The "*Baby Go*" application utilized two AI models for infant motor assessment during the study period. The first model was trained on standardized laboratory video data, while the second model was trained on both laboratory and home video data. The AI Model I showed high agreement between the parents' and physiotherapists' results (80%) and moderate agreement between physiotherapists' and AI results (70%). The AI Model II showed slightly reduced agreement: 76% for the parents' vs. physiotherapists' results, 66% for the physiotherapists' vs. AI results, and 61% for the parents' vs. AI results. These findings are consistent with the data of

doi:10.6342/NTU202504540

our earlier validation study, which demonstrated higher accuracy of AI classification for laboratory videos (0.91) compared to that for home videos (0.84) (Purwanto et al., 2025). It suggests that incorporating home video data into AI training may slightly reduce the model's performance in movement classification compared with expert assessments.

The results from AI Model I indicate a higher agreement between physiotherapist and parent assessments, suggesting that the training on laboratory samples effectively captured key developmental indicators. However, the introduction of AI Model II, which incorporated home video data, highlights the complexities of real-world applications, as evidenced by the slight decline in parent-physiotherapist agreement. This shift underscores the necessity for continuous refinement of AI models to ensure they remain relevant and accurate in diverse settings, as seen in other studies exploring technology's role in enhancing parental engagement and monitoring infant development (Achouche et al., 2024; Shorey et al., 2023). The discrepancy in assessment results between laboratory and home video data may stem from variations in environmental factors, such as lighting and background noise, which can influence the accuracy of movement recognition algorithms. Additionally, the controlled conditions of laboratory settings typically allow for more consistent and clear video quality, whereas home recordings can be more variable, impacting the overall performance of AI models in real-world scenarios. This highlights the need for ongoing advancements in AI technology to adapt to diverse environments and improve reliability in home-based assessments (Lin et al., 2024).

4.3 Technical Problems Affecting the AI Performance

Analysis of 936 home videos revealed that including all recordings regardless of orientation resulted in moderate agreement levels among parents, physiotherapists, and AI (77%, 67%, and 62%, respectively). However, when 498 well-oriented videos were analyzed

doi:10.6342/NTU202504540

separately, agreement rates improved to 79% between parents and physiotherapists, 79% between physiotherapists and AI, and 73% between parents and AI. These findings highlight the impact of video quality and orientation on assessment consistency, emphasizing the need for clearer guidance to parents during recording to support reliable AI-assisted evaluation of infant motor development. These findings highlight the critical impact of video quality and orientation on assessment consistency, emphasizing the need for clearer guidance to parents during recording to support reliable AI-assisted evaluation of infant motor development, as previously noted in systematic reviews of home-based assessment methods.

From a technical standpoint, correct camera orientation, stable lighting, and proper infant positioning are essential for accurately capturing subtle motor behaviors. Prior research has shown that home video-based assessments can be reliable when guided recording protocols are implemented (Kwong et al., 2022). In mobile health (mHealth) tools for developmental screening, inadequate video quality is a known source of misclassification and lower diagnostic concordance (Shorey et al., 2023). Accordingly, embedding real-time guidance features, such as prompts to rotate the camera to landscape mode, use stable backgrounds, and ensure full-body visibility, can improve video quality and assessment validity. In addition, the moderate performance of AI Model II, despite its machine learning base on more diverse home video data, indicates the limits of current algorithmic generalization. Variations in camera angle, length, camera resolution, and distance to the infants require additional processing steps before analysis (Passmore et al., 2024). Improving algorithm robustness may involve techniques such as video augmentation (e.g., simulating lighting and orientation changes), multi-angle training datasets, and the use of temporal sequence modeling (e.g., convolutional recurrent networks) to better capture dynamic movements (Zhou et al., 2017).

doi:10.6342/NTU202504540

4.4 Movement-based Assessment

The overall agreements in the movement-based assessment were moderate for the parents' vs. physiotherapists' results (79%), the physiotherapists' vs. the AI classification (79%), and the parents' vs. the AI results (73%). These results suggest that, although the AI model and parents were generally consistent with physiotherapist ratings, some variability remained in how movements were interpreted. This level of agreement highlights both the potential as well as the limitations of using AI-assisted tools and parental reports for remote assessment of infant motor development.

The overall agreement in this study was lower compared with our previous study, in which the accuracy with home videos reached 84% (Purwanto et al., 2025). This difference may be attributed to the larger and more diverse data used in the current study compared to the previous study. When compared with other AI-based applications, the results were broadly comparable. For example, the In-Motion with AI application classified about 80.2% of infant body key point positions correctly when compared with the General Movements Assessment (GMA) examiners' ratings (Adde et al., 2021). Similarly, the PTS-AI system, which was specifically developed for assessing the pull-to-sit movement, achieved average accuracies of 77.7% and 88.1% against expert evaluations using the Hammersmith Infant Neurological Examination (HINE) (Zhang et al., 2023). These findings suggest that the "*Baby Go*," which integrates AI with parental assessment, provides an accuracy level comparable to other AI-driven infant motor assessment tools.

An essential feature of the *"Baby Go"* application project was the evaluation of parental perception as a complementary method to AI-based analysis in the assessment of infant motor development. Parental reporting has been widely used in developmental screening, as parents are

doi:10.6342/NTU202504540

uniquely positioned to observe their children's behaviors across multiple contexts and time points (Libertus & Landa, 2013). However, the accuracy and consistency of parental judgments have been shown to vary substantially, influenced by both the characteristics of the observed behaviors and the caregivers' prior experiences and expectations (Bartlett, 1993). The findings in this study provide further evidence of these complexities, highlighting strengths and limitations inherent to parental assessments of infant motor assessment.

Movement-based agreement was lower in preterm infants compared with full-term infants, which can be explained by the tendency of preterm infants to show delayed and more variable motor development. Their motor behaviors are often less mature and less predictable, making it more challenging for parents to accurately recognize and assess these movements. In contrast, full-term infants typically demonstrate more stable and age-appropriate motor skills, facilitating higher parental agreement with professional assessments (Cuesta-Gómez et al., 2024).

Based on total agreement by position, parental perception results from the application showed moderate agreement with the physiotherapists' labeling results for the supine (67 %) and sitting movements (77%), and high agreement for the prone (91%) and standing movements (82 %). The agreement for the supine position was similar to that reported in an earlier study by Hsiao (2023), while the prone position showed a substantially increased agreement than in the earlier study. Improvements were also observed in the sitting and standing positions. These improvements are likely due to clearer movement descriptions and parental instructions in the application, in which parents were further supported with a written user guide and could receive assistance whenever they encountered technical difficulties using the application, either through email or phone calls. Similar to our approach, Lima et al. (2022) reported that parents received

doi:10.6342/NTU202504540

standard written instructions and guidance, as well as provided guidance messages for making the video recordings correctly, which contributed to more reliable remote home video assessments of infant motor assessment using AIMS recorded by mothers at home.

In the sub-analysis of the supine position, specific movements showed considerable disagreement between “Forearm Support (1)”, “Forearm Support (2)”, and “Extended Arm Support”. Some parents selected “Forearm Support (1)” while physiotherapists observed “Forearm Support (2)”, and others selected “Extended Arm Support “. This discrepancy did not stem from differing interpretations of classification criteria but was primarily due to the presence of more advanced movements typically observed in full-term infants. By the time of video assessment, parents often found it difficult to identify previous, less mature movements that had already been surpassed. This observation aligns with findings that full-term infants generally demonstrate more advanced gross motor development compared to preterm infants (Kwong et al., 2022). This finding also prompted a change in the application’s logic, so that if infants can perform a more advanced movement, any earlier movements that were not uploaded are automatically considered passed in the application.

Some difficulties observed by parents in this study involved movements that required reaching, such as “Reach from Forearm Support”, “Reaching from Extended Arm Support”, and “Reach with Rotation in Sitting”. The lower AI accuracy for these movements was primarily due to the infant’s hands often being occluded by the parent while providing stimulation. Another challenge involved the classification of highly similar infant movements, often referred to as confusing movements. Parental analysis showed varying agreement rates for movements such as “Forearm Support (1) and (2)”, “Reciprocal Creeping (1) and (2)”, “Sitting without Arm Support (1) and (2)”, “Supported Standing (2) and (3)”, and “Cruising without and Cruising with

doi:10.6342/NTU202504540

rotation". From the parental perspective, previous research indicates that recognition of specific postural characteristics often depends on the clarity of reference examples and training provided (Bosanquet et al., 2013). This limitation may lead parents to either overreport (assuming a more advanced milestone is achieved) or underreport (missing subtle transitions) certain movements. From the AI model's perspective, recent computer vision studies show that models trained on similar motor patterns often confuse movements when discriminative features are not distinct enough or are underrepresented in the training data (Zhang et al., 2020). This highlights the importance of expanding datasets to include more varied examples and detailed information about each movement, so the model can better recognize subtle differences in positions

Physiotherapists and the AI model generally showed high agreement for static movements, particularly in all supine positions, such as "Supine Lying (3)", "Supine Lying (4)" "Hands to Knees", and "Hands to Feet", static movements in prone positions, such as "Extended Arm Support" and static movements in sitting positions, such as "Supported Sitting" and "Sitting without Arm Support (1)." For dynamic movements, the agreement varied from moderate to low agreement depending on the complexity of the movements. This higher agreement for static movements may be due to their relative stability and predictability, which makes them easier for both humans and AI to recognize and assess consistently, compared with more complex dynamic movements (Ojeleye, 2023). Low agreement was persistently observed in reaching movements, such as "Reach from Forearm Support", "Reaching from Extended Arm Support" and "Reach with Rotation in Sitting", rotational movements, such as "Cruising with Rotation", and other confusing or highly similar movements across various positions, such as "Four-point Kneeling (1)", "Supported Standing (3)" "Pull to Stand with Support" "Pulls to Stand/Stands" and "Supported Standing with Rotation." These low accuracies indicate that the

doi:10.6342/NTU202504540

AI model had difficulty in classifying confusing movements in visually and kinematically similar fine-grained actions. This pattern is evident in the confusion matrix derived from laboratory video evaluations, underscoring the difficulty in distinguishing subtle developmental milestones based solely on pose sequences (Migliorelli et al., 2022). To enhance the accuracy of AI models in assessing dynamic movements and distinguishing between similar or confusing motor behaviors in infants, several strategies may be considered. Implementing a multi-stage data augmentation approach, which includes techniques such as random rotation, scaling, and occlusion simulation, may be explored in the future to improve model generalization by artificially increasing the diversity of training samples, particularly for underrepresented movements (Wang et al., 2020).

Considering on the movement-based agreement between parents, physiotherapists, and the AI model, adjustments may also be considered to improve the accuracy of this application in the future, we propose removing movements that persistently show low accuracy, such as reaching, including "Reach from Forearm Support," "Reaching from Extended Arm Support," and "Reach with Rotation in Sitting," and merging the most confusing movements into broader categories: "Forearm Support (1)" and "Forearm Support (2)" could be combined into "Forearm Support," "Reciprocal Creeping (1)" and "Reciprocal Creeping (2)" into "Reciprocal Creeping," "Sitting without Arm Support (1)" and "(2)" into "Sitting without Arm Support,", "Supported Standing (2)" and "(3)" into "Supported Standing", and "Cruising without rotation" and "Cruising with rotation" into "Cruising" movement. This simplification is feasible because, as indicated in the AIMS manual, these paired movements typically occur within closely overlapping age ranges and achievement windows, with credited criteria ranging between 50% and 90% (Piper & Darrah, 2021).

doi:10.6342/NTU202504540

Parental perception using the *"Baby Go"* application showed good accuracy compared to physiotherapist and AI assessments, although some movements were still challenging to identify precisely. Clinically, this means that combining parent perception with an AI model integrated in a mobile phone application can be a reliable tool for early infant motor screening, especially useful for remote or home-based monitoring where access to professionals is limited.

4.5 Age-based Assessment

In the age-based assessment of infants aged 3 to 18 months, agreement rates were 97% between parents and AI, 82% between physiotherapists and AI, and 74% between parents and physiotherapists. The results indicate generally good agreement among parental perception, the AI model, and physiotherapist evaluations, particularly within the 5 to 10-month age range. This pattern is consistent with prior research showing that the early acquisition of foundational gross motor movements in this age group, such as the development of sitting unsupported, reciprocal creeping, and supported standing, tends to be more easily recognized and reliably reported by both professionals and parents (Piek et al., 2008).

In contrast, lower agreement (below 60%) was observed at 4 and 11 months of age, with agreement reaching only 64% at 4 months due to disagreement in observing the movement "Forearm Support (1)," and just 50% at 11 months because of disagreement regarding "Sitting without Arm Support (2)." Early in infancy, motor skills are still emerging, and subtle or less differentiated movements (Adolph & Hoch, 2019) make it challenging for parents, AI, and professionals to classify movements consistently. Similarly, toward the end of the first year, individual variability in the timing and sequencing of motor milestones increases, influenced by differences in practice opportunities, environmental contexts, and biological maturation, complicating uniform classification (Piek et al., 2008).

doi:10.6342/NTU202504540

Additionally, the parental perception results demonstrate that parents can effectively assess infant motor development when supported by structured visual tools. We found similarly high agreement rates between parental and professional assessments for both full-term (95%) and preterm (95%) infants, suggesting that, with proper guidance, parents can reliably monitor development regardless of birth status. These results align with previous research showing that structured observation tools enhance parental reporting accuracy (Libertus & Landa, 2013). The validity of parental assessments in our study likely results from standardized visual demonstrations, clear developmental criteria, and opportunities for repeated assessments, supporting growing evidence that parental involvement has great potential to improve the accuracy and feasibility of developmental monitoring in infants, particularly those at high risk for neurodevelopmental delays (Marschik et al., 2023).

The validity of parental perception compared to physiotherapist assessments in this study was generally high across all infant age groups, with an overall agreement rate of 93%. This shows that parents were able to reliably recognize whether their infants demonstrated age-appropriate motor movements. Agreement rates were perfect (100%) in many age groups, including 3, 4, 5, 7, 10, 12, and 13–18 months, where parents and professionals fully agreed on whether development was within or not within the expected range. At 6 months, agreement remained strong (88%), although sensitivity decreased to 50%, indicating some cases of motor delay were missed by parents. Similarly, at 8 months, specificity decreased slightly to 88% and PPV to 80%, reflecting occasional over-identification of delays. These findings are in line with previous research showing that core gross motor milestones, such as sitting without support or creeping, tend to occur within predictable time frames and are frequently observed by parents in daily routines (Piek et al., 2008). However, the minor differences in sensitivity and specificity in

doi:10.6342/NTU202504540

some age groups highlight the importance of giving parents clear instructions and visual examples to help them recognize less obvious or transitional motor skills accurately (Boonzaaijer et al., 2019).

Four infants (28.6%) were classified as false positives, where the AI-parent assessment indicated motor delay but the AIMS scores fell within the normal range (25th to 90th percentile), suggesting some over-identification of motor delay with our AI model. However, there were no false negatives in this sample, meaning no infants were identified as normal by the AI-parent assessment but delayed by AIMS. The laboratory validation using the Alberta Infant Motor Scale (AIMS) demonstrated that most classifications based on parental perception were consistent with physiotherapy assessments. The validity analysis of AI results compared with onsite AIMS assessment showed an overall agreement of 79%, with a sensitivity of 100%, specificity of 75%, PPV of 40%, and NPV of 100%. In the sample, the majority of cases were true negatives, in which infants identified as being within the normal developmental range by parents and AI were also confirmed as normal on the AIMS assessment. This finding suggests that the system was generally reliable in ruling out motor delays. However, there were three false-positive cases, where parents classified infants as not within the normal range, while AIMS results indicated normal development. These discrepancies may reflect the tendency of some parents to be cautious or to overestimate subtle deviations from typical motor development. Similar patterns have been reported in other studies, where parents sometimes interpret minor variations in movement as signs of delay, especially when using structured checklists (Boonzaaijer et al., 2019). No false negatives were found in the validation, meaning that all infants with clear signs of delay were correctly identified by parents and confirmed by the AIMS assessment in the

doi:10.6342/NTU202504540

laboratory. This finding suggests that combining parental perception with screening tools can help identify infants needing further evaluation (Glascoe, 1997).

4.6 Clinical Implications

The "Baby Go" application allows parents to monitor their infant's motor development at home. AI classification can help parents review their observations and detect potential concerns regarding delayed motor development. Pediatricians and physiotherapists can use the application to track developmental progress and discuss results with families. The moderate agreement between parents' and physiotherapists' results, and between AI and physiotherapists' results with the current version of the application, indicates that both parental observations and AI analysis require cross-validation with physiotherapists or pediatricians, particularly for those infants classified as development not within the expected range.

The "*Baby Go"* application can clinically supplement parental monitoring for early screening of infant motor development, support pediatricians and physiotherapists in guiding early interventions, and promote accessibility in remote areas. In addition, the application can help parents better understand their infants' development. We recommend using the application while ensuring that parental observations and self-assessment have always followed the standardized assessments designed and monitored by health professionals.

4.7 Limitations and Future Study

This study has several limitations of note. First, the accuracy of the AI assessment was affected by technical problems, particularly issues with video recording and uploading in portrait mode and incorrect orientation, which reduced the AI's accuracy in both movement-based and age-based assessments. Second, the AI model was dependent on parental perception selection, which decreased classification accuracy. To address these issues, future work may need to

doi:10.6342/NTU202504540

mitigate these technical problems by revising the AI algorithm to provide automatical detection of orientation-error videos before movement recognition. In addition, future study may consider re-arranging the movements by removing or merging those with lower accuracy to enhance AI performance. Furthermore, integrating the 3D pose estimation technology may help enhance the AI accuracy. Finally, implementing real-time feedback in the application may help parents use the monitoring tools more effectively for improved motor developmental screening.

preprocessing orientation-error videos before AI movement recognition.

doi:10.6342/NTU202504540

Chapter 5. Conclusion

The upgraded "Baby Go" application successfully integrates AI-based assessment with parental education features, but requires further refinement of the parental perception feature. After exclusion of the videos with technical problems, parental perception and AI assessment demonstrate nearly high agreement for movement-based assessment and high agreement for age-based assessment. The parental perception and AI assessment results showed acceptable sensitivity, specificity, and negative predictive value, but low positive predictive value compared to the AIMS assessment. The findings indicate the supplemental use of the "Baby Go" application for parental monitoring in early screening of infant motor development, support pediatricians and physiotherapists in developmental follow-up and early interventions for full-term and preterm infants, but its parental observations and self-assessment have always followed the standardized assessments by health professionals.

doi:10.6342/NTU202504540

References

Abbott, A., Bartlett, D., Fanning, J., & Kramer, J. (2000). Infant Motor Development and Aspects of the Home Environment. *Pediatric Physical Therapy*, *12*(2), 62-67.

Achouche, V., Piollet, M., Temame, M., Cao, X. N., D'herbemont, V., Moreau, J., et al. (2024). The impact of a mobile application on parental attitudes, their knowledge of child development, and sense of parenting self-competence: A pilot study. *Annales Médico-Psychologiques, Revue Psychiatrique*, *182*(3), 233-241.

Adde, L., Brown, A., van den Broeck, C., DeCoen, K., Eriksen, B. K., Fjørtoft, T., et al. (2021). In-motion-app for remote general movement assessment: A multi-site observational study. *BMJ Open*, *11*(3), 1-9.

Adolph, K. E., & Hoch, J. E. (2019). Motor development: Embodied, embedded, enculturated, and enabling. *Annual Review of Psychology*, *70*(1), 141-164.

Agarwal, P. K., Xie, H., Rema, A. S. S., Rajadurai, V. S., Lim, S. B., Meaney, M., et al. (2020). Evaluation of the Ages and Stages Questionnaire (ASQ 3) as a developmental screener at 9, 18, and 24 months. *Early Human Development*, *147*, 1-7.

Ahn, H. Y., Ko, H. J., & Jo, H. J. (2023). Development and Effects of Mobile-Application-Based Parenting Support Program for Premature Infants' Mothers. *Healthcare*, *11*(19), 1-16.

Baccinelli, W., Bulgheroni, M., Simonetti, V., Fulceri, F., Caruso, A., Gila, L., et al. (2020). Movidea: A software package for automatic video analysis of movements in infants at risk for neurodevelopmental disorders. *Brain Sciences*, *10*(4), 203, 1-12.

Bartlett, D. J. (1993). *Validity of mothers' assessments of infant motor development.* Physical Therapy Department, University of Alberta. Edmonton, Alberta, Canada.

Bayley, N. (2006). *Bayley Scales of Infant and Toddler Development* (3rd Edition). Harcourt Assessment.San Antonio, USA.

Blencowe, H., Cousens, S., Chou, D., Oestergaard, M., Say, L., et al. (2013). Born Too Soon: The global epidemiology of 15 million preterm births. *Reproductive Health*, *10*(1), 1-14.

Boonzaaijer, M., van Dam, E., van Haastert, I. C., & Nuysink, J. (2017). Concurrent validity between live and home video observations using the Alberta Infant Motor Scale. *Pediatric Physical Therapy*, *29*(2), 146-151.

Boonzaaijer, M., van Wesel, F., Nuysink, J., Volman, M. J., & Jongmans, M. J. (2019). A home-video method to assess infant gross motor development: parent perspectives on feasibility. *BMC Pediatrics*, *19*(1), 1-12.

Bosanquet, M., Copeland, L., Ware, R., & Boyd, R. (2013). A systematic review of tests to predict cerebral palsy in young children. *Developmental Medicine & Child Neurology*, *55*(5), 418-426.

Bruininks, R. H., & Bruininks, B. D. (1978). *Bruininks-Oseretsky Test of Motor Proficiency*. American Guidance Service. Minnesota, USA.

doi:10.6342/NTU202504540

Burns, Y. R. (1992). *NSMDA: Physiotherapy Assessment for Infants & Young Children*. Copyright Publishing. Sidney, Australia.

Campbell, S. K., Swanlund, A., Smith, E., Liao, P.-j., & Zawacki, L. (2008). Validity of the TIMPSI for estimating concurrent performance on the test of infant motor performance. *Pediatric Physical Therapy:*, *20*(1), 3-10.

Campbell, S. K., Swanlund, A., Smith, E., Liao, P.J., & Zawacki, L. (2008). Validity of the TIMPSI for estimating concurrent performance on the test of infant motor performance. *Pediatric Physical Therapy:*, *20*(1), 3-10.

Cao, G., Liu, J., & Liu, M. (2022). Global, regional, and national incidence and mortality of neonatal preterm birth, 1990-2019. *JAMA Pediatrics*, *176*(8), 787-796.

Carreira, J., & Zisserman, A. (2017). Quo vadis, action recognition? A new model and the kinetics dataset. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA.

Carvajal, D. N., & Rowe, P. C. (2010). Research and statistics: Sensitivity, specificity, predictive values, and likelihood ratios. *Pediatrics in Review*, *31*(12), 511-513.

Chawanpaiboon, S., Vogel, J. P., Moller, A. B., Lumbiganon, P., Petzold, M., Hogan, D., et al. (2019). Global, regional, and national estimates of levels of preterm birth in 2014: a systematic review and modelling analysis. *The Lancet Global Health*, *7*(1), e37-e46. Check the name of of journal (Lancet Global Health?)

Chen, L.-C., Wu, Y.-C., Hsieh, W.-S., Hsu, C.-H., Leng, C.-H., Chen, W. J., et al., (2013). The effect of in-hospital developmental care on neonatal morbidity, growth and development of preterm Taiwanese infants: a randomized controlled trial. *Early Human Development*, *89*(5), 301-306.

Chen, H., Feng, R., Wu, S., Xu, H., Zhou, F., & Liu, Z. (2023). 2D Human pose estimation: A survey. *Multimedia Systems*, *29*(5), 3115-3138.

Chen, H. J., Ju Ko, M. H., Li, S. T., Chiu, N. C., & Hung, K. L. (2020). Prevalence of preschool children developmental disabilities in northeastern Taiwan - Screening with Taipei City Developmental Screening Checklist for Preschoolers, 2nd Version. *Journal of the Formosan Medical Association*, *119*(7), 1174-1179.

Cuesta-Gómez, A., Fernández-González, P., Carratalá-Tejada, M., & Aguilar-Bejines, I. (2024). Differences in Motor Development between Preterm Infants and Full-Term Preschool Children. *Children*, *11*(2), 1-10.

DeWitt, A., Kientz, J., & Liljenquist, K. (2022). Quality of mobile apps for child development support: search in app stores and content analysis. *Journal of Medical Internet Research Pediatrics and Parenting*, *5*(4), 1-14. What is the full name of this journal?

Duan, H., Zhao, Y., Chen, K., Lin, D., & Dai, B. (2022). *Revisiting skeleton-based action recognition* Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, New Orleans, LA, USA.

doi:10.6342/NTU202504540

Fish, L. A., & Jones, E. J. (2021). A survey on the attitudes of parents with young children on in-home monitoring technologies and study designs for infant research. *PLoS One*, *16*(2), 1-29.

Folio, M. R. (1983). Peabody Developmental Motor Scales. *DLM Teaching Resources*. Texas, USA.

Frankenburg, W. K., Dodds, J., Archer, P., Shapiro, H., & Bresnick, B. (1992). The Denver II: a major revision and restandardization of the Denver Developmental Screening Test. *Pediatrics*, *89*(1), 91-97.

Frey, H. A., & Klebanoff, M. A. (2016). The epidemiology, etiology, and costs of preterm birth. *Seminars in Fetal and Neonatal Medicine*, *21*(2), 68-73.

Glascoe, F. P. (1997). Parents' concerns about children's development: prescreening technique or screening test? *Pediatrics*, *99*(4), 522-528.

Grant, C. M., Jones, A., Crowe, M., & Land, H. (2023). Gross motor assessment tools for children 0-12 years of age using telehealth: a scoping review. *Internet Journal of Allied Health Sciences and Practice*, *21* (4), 1-19.

Griffiths, A., Toovey, R., Morgan, P. E., & Spittle, A. J. (2018). Psychometric properties of gross motor assessment tools for children: a systematic review. *BMJ Open*, *8*(10), 1-14.

Groos, D., Adde, L., Aubert, S., Boswell, L., De Regnier, R.-A., Fjørtoft, T., et al. (2022). Development and validation of a deep learning method to predict cerebral palsy from spontaneous movements in infants at high risk. *JAMA Netwoks Open*, *5*, e2221325-e2221325.

Henderson, S. E., Sugden, D., & Barnett, A. L. (1992). *Movement Assessment Battery for Children-2*. Pearson Assessment. London, UK.

Hernández, M. C., Canals, S. J., & Fernández, B., J. (2011). Parents' perceptions of their neonates and their relation to infant development. *Child: Care, Health and Development*, *37*(4), 484-492.

Hess, C. R., Papas, M. A., & Black, M. M. (2004). Use of the Bayley Infant Neurodevelopmental Screener with an environmental risk group. *Journal of Pediatric Psychology*, *29*(5), 321-330.

Hsiao, Y. C. (2023). *Application of Artificial Intelligence in Mobile Application for Infant Movement Assessment* School and Graduate Institute of Physical Therapy, National Taiwan University. Taipei, Taiwan.

Jeng, S.-F., Yau, K.-I. T., & Teng, R.-J. (1998). Neurobehavioral development at term in very low-birthweight infants and normal term infants in Taiwan. *Early Human Development*, *51*(3), 235-245.

Jeng, S. F., Yau, K. I. T., Chen, L. C., & Hsiao, S. F. (2000). Alberta infant motor scale: Reliability and validity when used on preterm infants in Taiwan. *Physical Therapy*, *80*, 168-178.

Jeng, S.-F., Lau, T.-W., Hsieh, W.-S., Luo, H.-J., Chen, P.-S., Lin, K.-H., et al., (2008). Development of walking in preterm and term infants: age of onset, qualitative features and sensitivity to resonance. *Gait &Posture*, *27*(2), 340-346.

doi:10.6342/NTU202504540

Kwong, A. K., Doyle, L. W., Olsen, J. E., Eeles, A. L., Zannino, D., Mainzer, R. M., et al. (2022). Parent-recorded videos of infant spontaneous movement: Comparisons at 3–4 months and relationships with 2-year developmental outcomes in extremely preterm, extremely low birthweight and term-born infants. *Paediatric and Perinatal Epidemiology*, *36*(5), 673-682.

Kwong, A. K., Eeles, A. L., Olsen, J. E., Cheong, J. L., Doyle, L. W., & Spittle, A. J. (2019). The baby moves smartphone app for general movements assessment: Engagement amongst extremely preterm and term-born infants in a state-wide geographical study. *Journal of Paediatrics and Child Health*, *55*(5), 548-554.

Lanzagorta-Ortega, D., Carrillo-Pérez, D., & Carrillo-Esper, R. (2022). Artificial intelligence in medicine: present and future. *Gaceta Medica de Mexico*, *158*(Suplement 1), 17-21.

Libertus, K., & Landa, R. J. (2013). The Early Motor Questionnaire (EMQ): A parental report measure of early motor development. *Infant Behavior and Development*, *36*(4), 833-842.

Lima, C. R. G., Dos Santos, A. N., Dos Santos, M. M., Morgan, C., & Rocha, N. A. C. F. (2022). Tele-care intervention performed by parents involving specific task- environment- participation (STEP protocol) for infants at risk for developmental delay: protocol of randomized controlled clinical trial. *BMC Pediatrics*, *22*(1), 51.

Lima, M., & Mondardini, M. C. (2023). Premature Infants. In *Frailty in Children: From the Perioperative Management to the Multidisciplinary Approach* (pp. 11-32). Springer, Cham, Switzerland.

Lin, S. C., Chandra, E., Tsao, P. N., Liao, W. C., Chen, W. J., Yen, T. A., et al. (2024). Application of artificial intelligence in infant movement classification: A reliability and validity study in infants who were full-term and preterm. *Physical Therapy*, *104*(2), 1-10.

Lipkin, P. H., Macias, M. M., Norwood, K. W., Brei, T. J., Davidson, L. F., Davis, B. E., et al. (2020). Promoting optimal development: identifying infants and young children with developmental disorders through developmental surveillance and screening. *Pediatrics*, *145*(1), 2019-3449.

Marschik, P. B., Kwong, A. K. L., Silva, N., Olsen, J. E., Rüther, M. S., Bölte, S., et al. (2023). Mobile solutions for clinical surveillance and evaluation in infancy-general movement apps. *Clinical Medicine*, *12*(10), 1-12.

McCarron, L. T. (1982). *MAND: McCarron Assessment of Neuromuscular Development, Fine and Gross Motor Abilities*. McCarron-Dial Systems. Dallas, Texas, USA.

Mendonça, B., Sargent, B., & Fetters, L. (2016). Cross-cultural validity of standardized motor development screening and assessment tools: A systematic review. *Developmental Medicine and Child Neurology, 58*(12), 1213-1222.

Migliorelli, L., Frontoni, E., & Moccia, S. (2022). An accurate estimation of preterm infants' limb pose from depth images using deep neural networks with densely connected atrous spatial convolutions. *Expert Systems with Applications*, *204*, 1-11.

Moccia, S., Migliorelli, L., Carnielli, V., & Frontoni, E. (2020). Preterm infants' pose estimation with spatio-temporal features. *IEEE Trans Biomed Engineering*, *67*(8), 2370-2380.

doi:10.6342/NTU202504540

Mu, S.-C., Tsou, K.-S., Hsu, C.-H., Fang, L.-J., Jeng, S.-F., Chang, C.-H., et al., (2008). Cognitive development at age 8 years in very low birth weight children in Taiwan. *Journal of the Formosan Medical Association*, *107*(12), 915-920.

Noritz, G. H., Murphy, N. A., Panel, N. S. E., Murphy, N. A., Hagan Jr, J. F., Lipkin, P. H., et al. (2013). Motor delays: early identification and evaluation. *Pediatrics*, *131*(6), e2016-e2027.

Ohuma, E. O., Moller, A.-B., Bradley, E., Chakwera, S., Hussain-Alkhateeb, L., Lewin, A., et al. (2023). National, regional, and global estimates of preterm birth in 2020, with trends from 2010: a systematic analysis. *The Lancet*, *402*(10409), 1261-1271.

Ojeleye, J. O. (2023). *Developing tools and methods for human body pose estimation* Computer Science, University of Eastern Finland. Finland.

Oudgenoeg-Paz, O., Atun-Einy, O., & van Schaik, S. D. M. (2020). Two Cultural Models on Infant Motor Development: Middle Class Parents in Israel and the Netherlands. *Frontiers in Psychology*, *11*, 119.

Piek, Dawson, L., Smith, L. M., & Gasson, N. (2008). The role of early fine and gross motor development on later motor and cognitive ability. *Human Movement Science 27*(5), 668-681.

Passmore, E., Kwong, A., Greenstein, S., Olsen, J., Eeles, A., Cheong, J., et al. (2024). Automated identification of abnormal infant movements from smart phone videos. *PLOS Digital Health*, *3*(2), 1-21.

Piper, M., & Darrah, J. (1994). *Alberta Infant Motor Scale* (1st Edition). WB Saunders, Philadelphia, PA, USA.

Piper, M., & Darrah, J. (2021). *Motor Assessment of the Developing Infant-E-Book: Alberta Infant Motor Scale (AIMS)* (2nd Edition). Elsevier Health Sciences, Philadelphia, PA, USA.

Piper, M. C., Pinnell, L. E., Darrah, J., Maguire, T., & Byrne, P. J. (1992). Construction and validation of the Alberta Infant Motor Scale (AIMS). *Canadian Journal of Public Health*, *83*, S46-50.

Pugliese, M., Rossi, C., Guidotti, I., Gallo, C., Della Casa, E., Bertoncelli, N., et al. (2013). Preterm birth and developmental problems in infancy and preschool age Part II: cognitive, neuropsychological and behavioural outcomes. *The Journal of Maternal-Fetal Neonatal Medicine*, *26*(16), 1653-1657.

Purwanto, Y., Chandra, E., Tsao, P.-N., Yen, T.-A., Liao, W.-C., Chen, W. J., et al. (2025). Artificial intelligence algorithms effectively classify 38 movements in infants born full-term and preterm recorded in the laboratory and at home. *Journal of the Formosan Medical Association*, (Available online 14 August 2025).

Purwanto, Y., Jeng, S.-F., Hsu, Y.-J., Tsao, P.-N., Yen, T. A., Liao, W. C., et al. (2023). Concurrent validity of the Alberta Infant Motor Scale – Short Version compared with the original version in term and preterm infants. The World Physiotherapy Congress, Dubai, United Arab Emirates.

Raina, V., Molchanova, N., Graziani, M., Malinin, A., Muller, H., Cuadra, M. B., et al. (2023). *Tackling bias in the dice similarity coefficient: Introducing nDSC for white matter lesion*

doi:10.6342/NTU202504540

segmentation 2023 IEEE 20th International Symposium on Biomedical Imaging (ISBI), Cartagena de Indias, Colombia.

Reich, S., Zhang, D., Kulvicius, T., Bölte, S., Nielsen-Saines, K., Pokorny, F. B., et al. (2021). Novel AI-driven approach to classify infant motor functions. *Scientific Reports*, *11*(1), 1-13.

Rezaei, B., Christakis, Y., Ho, B., Thomas, K., Erb, K., Ostadabbas, S., et al. (2019). Target-specific action classification for automated assessment of human motor behavior from video. *Sensors (Basel)*, *19*(19), 1-18.

Rosales, M. R., Simsic, J., Kneeland, T., & Heathcock, J. (2024). Quantifying Arm and Leg Movements in 3-Month-Old Infants Using Pose Estimation: Proof of Concept. *Sensors*, *24*(23), 1-14.

Rosenbaum, D. A., Carlson, R. A., & Gilmore, R. O. (2001). Acquisition of intellectual and perceptual-motor skills. *Annual Review of Psychology*, *52*(1), 453-470.

Shi, L., Zhang, Y., Cheng, J., & Lu, H. (2021). AdaSGN: Adapting joint number and model size for efficient skeleton-based action recognition. Proceedings of the IEEE/CVF International Conference on Computer Vision, Beijing, China.

Shorey, S., Chong, Y. S., Shi, L., Chua, J. S., Mathews, J., Lim, S. H., et al. (2023). Evaluating the effects of the supportive parenting App on infant developmental outcomes: longitudinal study. *Journal of Medical Internet Research mHealth and uHealth*, *11*(1), 1-16.

Shrivastava, R., & Pandey, M. (2017). Human activity recognition by analysis of skeleton joint position in internet of things (IOT) environment. *Indian Journal of Science and Technology*, *10*(16), 1-9.

Silva, S. D., Flôres, F. S., Corrêa, S. L., Cordovil, R., & Copetti, F. (2017). Mother's perception of children's motor development in southern Brazil. *Perceptual Motor Skills*, *124*(1), 72-85.

Šimundić, A. M. (2009). Measures of diagnostic accuracy: basic definitions. *European Journal of the International Federation of Clinical Chemistry and Laboratory Medicine. 19*(4), 203-211.

Spittle, A. J., Olsen, J., Kwong, A., Doyle, L. W., Marschik, P. B., Einspieler, C., et al. (2016). The Baby Moves prospective cohort study protocol: using a smartphone application with the General Movements Assessment to predict neurodevelopmental outcomes at age 2 years for extremely preterm or extremely low birthweight infants. *BMJ Open*, *6*(10), 1-8.

Su, Y. H., Jeng, S. F., Hsieh, W. S., Tu, Y. K., Wu, Y. T., & Chen, L. C. (2017). Gross motor trajectories during the first year of life for preterm infants with very low birth weight. *Physical Therapy*, *97*(3), 365-373.

Suir, I., Oosterhaven, J., Boonzaaijer, M., Nuysink, J., & Jongmans, M. (2022). The AIMS home-video method: parental experiences and appraisal for use in neonatal follow-up clinics. *BMC Pediatrics*, *22*(1), 1-11.

Svensson, K. A., Örtqvist, M., Bos, A. F., Eliasson, A. C., & Sundelin, H. E. (2021). Usability and inter-rater reliability of the neuromotion app: A tool in general movements assessments. *European Journal of Paediatric Neurology*, *33*, 29-35.

doi:10.6342/NTU202504540

Syrengelas, D., Siahanidou, T., Kourlaba, G., Kleisiouni, P., Bakoula, C., & Chrousos, G. (2010). Standardization of the Alberta infant motor scale in full-term Greek infants: Preliminary results. *Early Human Development*, *86*(4), 245-249.

Tieman, B. L., Palisano, R. J., & Sutlive, A. C. (2005). Assessment of motor development and function in preschool children. *Developmental Disabilities Research Reviews*, *11*(3), 189–196.

Tsai, Y.-H., Purwanto, Y., Chandra, E., Hsieh, C. W., Tsao, P.-N., Yen, T.-A., et al. (2025). A Feasibility Study of a Video-Based Application by Parents of Infants Born Full-Term and Preterm. *Pediatric Physical Therapy*, *37*(3), 317-325.

Uesugi, M., Tokuhisa, K., & Shimada, T. (2008). The reliability and validity of the Alberta Infant Motor Scale in Japan. *Journal of Physical Therapy Science*, *20*, 169-175.

Ulrich, D. A. (1985). Test of gross motor development. *Journal of Motor Learning and Development*, *17*(3), 247–258.

Valentini, N. C., & Saccani, R. (2012). Brazilian validation of the Alberta Infant Motor Scale. *Physical Therapy*, *92*(3), 440-447.

van Iersel, P. A., la Bastide-van Gemert, S., Wu, Y.-C., & Hadders-Algra, M. (2020). Alberta Infant Motor Scale: Cross-cultural analysis of gross motor development in Dutch and Canadian infants and introduction of Dutch norms. *Early Human Development*, 151, 1-5.

Wang, H., Li, H., Wang, J., & Jin, H. (2018). Reliability and concurrent validity of a Chinese version of the Alberta Infant Motor Scale administered to high-risk infants in China. *BioMed Research International*, *5*, 1-10.

Wang, J., Sun, K., Cheng, T., Jiang, B., Deng, C., Zhao, Y., et al. (2021). Deep High-Resolution Representation Learning for Visual Recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *43*(10), 3349-3364.

Wang, Y. (2022). *Cultural differences in infant motor development: A comparison of early locomotor experience,* Developmental Psychology Department, Simon Fraser University. British Columbia, Canada.

Wang, Y., Yao, Q., Kwok, J. T., & Ni, L. M. (2020). Generalizing from a few examples: A survey on few-shot learning. *ACM Computing Surveys*, *53*(3), 1-34.

Woodward, L. J., Moor, S., Hood, K. M., Champion, P. R., Foster-Cohen, S., Inder, T. E., et al. (2009). Very preterm children show impairments across multiple neurodevelopmental domains by age 4 years. *Archives of Disease in Childhood-Fetal Neonatal Edition*(5), F339-344.

Yao, B., Jiang, X., Khosla, A., Lin, A. L., Guibas, L., & Fei, L. F. (2011, 6-13 Nov. 2011). Human action recognition by learning bases of action attributes and parts. Proceedings of the IEEE International Conference on Computer Vision, Barcelona, Spain.

Yeh, K. K., Liu, W. Y., Wong, A. M. K., & Lein, R. (2020). Validity of General Movement Assessment Based on Clinical and Home Videos. *Pediatric Physical Therapy*, *32*(1), 35-43.

doi:10.6342/NTU202504540

Yildirim, C., Asalioğlu, A., Coşkun, Y., Acar, G., & Akman, İ. (2022). General movements assessment and Alberta Infant Motor Scale in neurodevelopmental outcome of preterm infants. *Pediatrics and Neonatology*, *63*(5), 535-541.

Yu, Y.-T., Hsieh, W.-S., Hsu, C.-H., Lin, Y.-J., Lin, C.-H., Hsieh, S., Lu, L., et al., (2017). Family-centered care improved neonatal medical and neurobehavioral outcomes in preterm infants: randomized controlled trial. *Physical Therapy*, *97*(12), 1158-1168.

Zhang, J., Lim, J., Kim, M.-H., Hur, S., & Chung, T.-M. (2023). WM–STGCN: a novel spatiotemporal modeling method for Parkinsonian gait recognition. *Sensors*, *23*(10), 1-20.

Zhang, P., Lan, C., Zeng, W., Xing, J., Xue, J., & Zheng, N. (2020). Semantics-guided neural networks for efficient skeleton-based human action recognition. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA.

Zhou, X., Huang, Q., Sun, X., Xue, X., & Wei, Y. (2017). Weakly-supervised transfer for 3d human pose estimation in the wild. IEEE international conference on computer vision, ICCV, Venice, Italy.

鄭, 素. 芳., & 楊, 佩. 瑜. (2021). *與巴掌仙子共舞-週歲前的育兒技巧. 臺北臺灣*. 金名圖書公司.

doi:10.6342/NTU202504540

Tables and Figures

Table 1-1. Infants' perinatal data (Cut-off 2025-06-30)

Characteristics	**Preterm (N=29)**	**Full-Term (N=16)**	**All (N=45)**
Female sex	16 (55.2%)	10 (62.5%)	26 (57.8%)
Birth body weight (grams)	1,756 ± 362.6*	2,949 ± 269.9*	2,353 ± 319.7
Gestational age (weeks)	33.4 ± 2.1*	38.6 ± 0.8*	36.0 ± 3.7
Apgar score at 1 min	6.8 ± 2.2*	8.2 ± 1.5*	7.5 ± 1.0
Apgar score at 5 min	8.3 ± 1.3*	8.9 ± 0.3*	8.6 ± 0.4
Small for gestational age	4 (13.8%)	2 (12.5%)	6 (13.3%)
Respiratory distress syndrome			
Normal	7 (24.1%)*	16 (100%)*	23 (51.1%)
Grade 1	16 (55.2%)*	0 (0.0%) *	16 (35.6%)
Grade 2	4 (13.8%)*	0 (0.0%)*	4 (8.9%)
Grade 3	2 (6.9%)*	0 (0.0%)*	2 (4.4%)
Bronchopulmonary dysplasia			
Normal	20 (69.0%)*	16 (100%)*	36 (80.0%)
Mild	4 (13.8%)*	0 (0.0%)*	4 (8.9%)
Moderate	5 (17.2%)*	0 (0.0%)*	5 (11.1%)
Severe	0 (0.0%)*	0 (0.0%)*	0 (0.0%)
Intraventricular hemorrhage			
Normal	28 (96.6%)	16 (100%)	44 (97.8%)
Grade I-II	1 (3.4%)	0 (0%)	1 (2%)
Grade II-IV	0 (0%)	0 (0%)	0 (0%)
Periventricular Leukomalacia	0 (0%)	0 (0%)	0 (0%)
Retinopathy of Prematurity	0 (0%)	0 (0%)	0 (0%)
Patent ductus arteriosus requiring ligation	0 (0%)	0 (0%)	0 (0%)
Necrotizing enterocolitis	0 (0%)	0 (0%)	0 (0%)
Hyperbilirubinemia requiring phototherapy	3 (10.4%)	0 (0.0%)	3 (6.7%)

Data presented as mean ± SD or number (%).

* $P<0.05$.

doi:10.6342/NTU202504540

Table 1-2. Parents' sociodemographic data (Cut-off 2025-06-30)

Characteristics	Preterm (N=28)	Full-term (N=16)	All (N=45)
Paternal age (years)	41.0 ± 7.2*	37.7 ± 3.6*	39.7 ± 5.6
Paternal education			
College or above	28 (96.6%)	15 (93.8%)	43 (95.6%)
Senior high school	1 (3.4%)	1 (6.2%)	2 (4.4%)
Middle school or below	0 (0.0%)	0 (0.0%)	0 (0%)
Paternal occupation			
Professional	22 (75.9%)	14 (87.5%)	36 (80.0%)
Technological	5 (17.2%)	2 (12.5%)	7 (15.6%)
Labor or homemaker	2 (6.9%)	0 (0.0%)	2 (4.4%)
Maternal age (years)	38.5 ± 4.3	35.8 ± 3.5	37.2 ± 1.9
Maternal education			
College or above	27 (93.1%)	16 (100.0%)	43 (95.6%)
Senior high school	2 (6.9%)	0 (0.0%)	2 (4.4%)
Middle school or below	0 (0.0%)	0 (0.0%)	0 (0%)
Maternal occupation			
Professional	17 (58.6%)	10 (62.4%)	27 (60.0%)
Technological	3 (10.3%)	3 (18.8%)	6 (13.3%)
Labor or homemaker	9 (31.1%)	3 (18.8%)	12 (26.7%)
Annual household income (NTD)			
> 1,500,000	10 (34.5%)	10 (62.5%)	20 (44.4%)
1,000,000-1,500,000	11 (37.9%)	4 (25.0%)	15 (33.3%)
< 1,000,000	8 (27.6%)	2 (12.5%)	10 (22.2%)
Parent as caregiver	25 (86.2%)	12 (75.0%)	37 (82.2%)
Full-time maternal caregiving	23 (79.3%)*	8 (50.0%)*	31 (68.9%)
Family size	4.9 ± 1.7*	3.6 ± 1.0*	4.3 ± 0.9
Number of children	2.0 ± 0.7*	1.4 ± 0.5*	1.7 ± 0.4
First child	12 (41.4%)	10 (62.5%)	22 (48.9%)

Data presented as mean ± SD or number (%).
One father of a twin in the preterm group had passed away during the mother's pregnancy.
NTD=New Taiwan dollars.
* P<0.05.

doi:10.6342/NTU202504540

Table 2. Number of videos uploaded and uploaded trials by preterm and full-term infants for movement- and age-based assessment (Data from Application, Cut off, 2025-06-30)

Age (months)	Number of Videos Uploaded for Movement-Based Assessment			Number of Uploading Trials for Age-Based Assessment		
	Preterm	Full-term	Total	Preterm	Full-term	Total
3	31 6%)	11 (2.8%)	42	8 (6.2%)	4 (4.5%)	12
4	50 (5.7%)	32 (8.1%)	82	14 (10.9%)	7 (7.9%)	21
5	94 (17.3%)	90 (22.9%)	184	18 (14.0%)	11 (12.4%)	29
6	65 (12.0%)	56 (14.2%)	121	18 (14.0%)	14 (15.7%)	32
7	67 (12.3%)	63 (16.0%)	130	16 (12.4%)	14 (15.7%)	30
8	73 (13.4%)	40 (10.2%)	113	10 (7.8%)	12 (13.5%)	22
9	34 (7.5%)	49 (12.5%)	83	12 (9.3%)	9 (10.1%)	21
10	83 (15.3%)	28 (7.1%)	111	10 (7.8%)	9 (10.1%)	19
11	25 (5.5%)	20 (5.1%)	45	8 (6.2%)	6 (6.7%)	14
12	10 (1.8%)	2 (0.5%)	12	5 (3.9%)	2 (2.2%)	7
13	2 (0.4%)	0 (0.0%)	2	4 (3.1%)	1 (1.1%)	5
14	1 (0.2%)	2 (0.5%)	3	3 (2.3%)	0 (0.0%)	3
15	2 (0.4%)	0 (0.0%)	2	2 (1.6%)	0 (0.0%)	2
16	6 (1.1%)	0 (0.0%)	6	1 (0.8%)	0 (0.0%)	1
17	0 (0%)	0 (0.0%)	0	0 (0.0%)	0 (0.0%)	0
18	0 (0%)	0 (0.0%)	0	0 (0.0%)	0 (0.0%)	0
Total	**543 (58%)**	**393 (42%)**	**936**	**129 (59%)**	**89 (41%)**	**218**

doi:10.6342/NTU202504540

Table 3. Number of videos, complete and incomplete set of movements uploaded by parents of preterm and full-term infants
(Data from Application, Cut off, 2025-06-30)

Age (months)	Preterm					Full-term				
	Complete		Incomplete	No upload	Total	Complete		Incomplete	No upload	Total
	Pass	Not pass				Pass	Not pass			
3	3 (38%)	1 (13%)	4 (50%)	1 (13%)	8	0 (0%)	3 (75%)	0 (0%)	1 (25%)	4
4	7 (50%)	2 (14%)	5 (36%)	4 (29%)	14	2 (29%)	2 (29%)	3 (43%)	0 (0%)	7
5	5 (28%)	2 (11%)	10 (56%)	4 (22%)	18	6 (55%)	1 (9%)	3 (27%)	1 (9%)	11
6	3 (17%)	1 (6%)	10 (56%)	3 (17%)	18	5 (36%)	0 (0%)	7 (50%)	2 (14%)	14
7	1 (6%)	4 (25%)	7 (44%)	4 (25%)	16	6 (43%)	1 (7%)	5 (36%)	2 (14%)	14
8	0 (0%)	2 (20%)	5 (50%)	3 (30%)	10	3 (25%)	2 (17%)	5 (42%)	2 (17%)	12
9	1 (8%)	1 (8%)	6 (50%)	3 (25%)	12	3 (33%)	1 (11%)	3 (33%)	2 (22%)	9
10	2 (20%)	1 (10%)	4 (40%)	3 (30%)	10	2 (22%)	2 (22%)	2 (22%)	3 (33%)	9
11	1 (13%)	3 (38%)	1 (13%)	2 (25%)	8	0 (0%)	0 (0%)	3 (50%)	3 (50%)	6
12	1 (20%)	0 (0%)	1 (20%)	3 (60%)	5	1 (50%)	0 (0%)	1 (50%)	0 (0%)	2
13	0 (0%)	0 (0%)	2 (50%)	2 (50%)	4	0 (0%)	0 (0%)	1 (100%)	0 (0%)	1
14	0 (0%)	0 (0%)	0 (0%)	1 (33%)	3	0 (0%)	0 (0%)	0 (0%)	0 (0%)	**0**
15	0 (0%)	0 (0%)	0 (0%)	0 (0%)	2	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0
16	0 (0%)	0 (0%)	0 (0%)	0 (0%)	1	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0
17	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0
18	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0
Total	**24 (19%)**	**17 (13%)**	**55 (43%)**	**33 (26%)**	**129**	**28 (31%)**	**12 (13%)**	**16 (18%)**	**33 (37%)**	**89**

doi:10.6342/NTU202504540

Table 4. Summary of the agreement for movement-based assessment results across application versions and AI models

Application Conditions	Time Period	Major Changes	No. of Data Set	Agreement		
				PT vs. Parent	PT vs. AI	Parent vs. AI
Versions						
Version 2.0	2024-10-04 to 2024-12-5	Addition of assessment, follow-up, and education features and reminders, and a video editor	78	76%	74%	68%
Version 2.1	2024-12-05 to 2025-02-18	Changes in the uploading interface and wording revision for recording time duration	199	82%	65%	60%
Version 2.2	2025-02-18 to 2025-03-08	Changes in the logic for advanced movements and added guidelines to identify confusing movements.	39	76%	85%	74%
Version 2.3	2025-03-08 to 2025-06-30	Accommodation for the vertical position problem during uploading videos	620	75%	65%	60%
AI Models						
AI Model I	2024-10-04 to 2025-01-14	The application applied the AI model trained on 11,454 laboratory video samples	194	80%	70%	64%
AI Model II	2025-01-14 to 2025-06-30	The application applied the AI model trained on the 11,454 laboratory and 2,054 home video samples	742	76%	66%	61%
All Data	2025-01-14 to 2025-06-30	Portrait and incorrect orientation videos were included	936	77%	67%	62%
All Data	2025-01-14 to 2025-06-30	Portrait and incorrect orientation videos were excluded	498	79%	79%	73%

PT vs. Parent = Agreement between the physiotherapists' and parents' assessment results.
PT vs. AI = Agreement between the physiotherapists' and AI classification results.
Parent vs. AI = Agreement between the parents' and AI classification results.

doi:10.6342/NTU202504540

Table 5. Agreement of movement-based assessment results for application version 2.0 (Cut-off 2024-12-05)

Movements	No. uploaded	Parent vs. PT		AI vs. PT		Parent vs. AI	
		Agreed	Disagreed	Agreed	Disagreed	Agreed	Disagreed
Forearm support (1)	16	8 (50%)	8 (50%)	11 (69%)	5 (31%)	8 (50%)	16 (50%)
Forearm support (2)	1	1 (100%)	0 (0%)	1 (100%)	0 (0%)	1 (100%)	1 (0%)
Extend arm support	8	4 (50%)	4 (50%)	5 (63%)	3 (38%)	2 (25%)	8 (75%)
Reach from forearm support	1	1 (100%)	0 (0%)	1 (100%)	0 (0%)	1 (100%)	1 (0%)
Pivoting	1	1 (100%)	0 (0%)	1 (100%)	0 (0%)	1 (100%)	1 (0%)
Four-point kneeling (1)	1	1 (100%)	0 (0%)	0 (0%)	1 (100%)	0 (0%)	1 (100%)
Reciprocal crawling	1	1 (100%)	0 (0%)	1 (100%)	0 (0%)	1 (100%)	1 (0%)
Four-point kneeling to sitting	1	1 (100%)	0 (0%)	1 (100%)	0 (0%)	1 (100%)	1 (0%)
Reciprocal creeping (1)	1	1 (100%)	0 (0%)	1 (100%)	0 (0%)	1 (100%)	1 (0%)
Reaching from extended arm support	1	1 (100%)	0 (0%)	0 (0%)	1 (100%)	0 (0%)	1 (100%)
Reciprocal creeping (2)	1	1 (100%)	0 (0%)	1 (100%)	0 (0%)	1 (100%)	1 (0%)
Supine lying (3)	10	9 (90%)	1 (10%)	8 (80%)	2 (20%)	8 (80%)	10 (20%)
Supine lying (4)	6	3 (50%)	3 (50%)	5 (83%)	1 (17%)	5 (83%)	6 (17%)
Hands to knees	2	2 (100%)	0 (0%)	2 (100%)	0 (0%)	2 (100%)	2 (0%)
Hands to feet	4	4 (100%)	0 (0%)	4 (100%)	0 (0%)	4 (100%)	4 (0%)
Rolling supine to prone with rotation	2	2 (100%)	0 (0%)	2 (100%)	0 (0%)	2 (100%)	2 (0%)
Sitting with support	9	8 (89%)	1 (11%)	4 (44%)	5 (56%)	5 (56%)	9 (44%)
Sitting with propped arms	2	1 (50%)	1 (50%)	2 (100%)	0 (0%)	1 (50%)	2 (50%)
Pull to sit	1	1 (100%)	0 (0%)	0 (0%)	1 (100%)	0 (0%)	1 (100%)
Sitting with arm support	1	1 (100%)	0 (0%)	1 (100%)	0 (0%)	1 (100%)	1 (0%)
Sitting without arm support (1)	1	0 (0%)	1 (100%)	0 (0%)	1 (100%)	1 (100%)	1 (0%)
Reach with rotation in sitting	0	-	-	-	-	-	-
Sitting to four-point kneeling	1	1 (100%)	0 (0%)	1 (100%)	0 (0%)	1 (100%)	1 (0%)
Sitting without arm support (2)	1	1 (100%)	0 (0%)	1 (100%)	0 (0%)	1 (100%)	1 (0%)
Supported standing (2)	1	1 (100%)	0 (0%)	1 (100%)	0 (0%)	1 (100%)	1 (0%)
Supported standing (3)	1	1 (100%)	0 (0%)	1 (100%)	0 (0%)	1 (100%)	1 (0%)
Pull to stand with support	0	-	-	-	-	-	-
Pulls to stand/stands	1	1 (100%)	0 (0%)	1 (100%)	0 (0%)	1 (100%)	1 (0%)
Supported standing with rotation	1	1 (100%)	0 (0%)	1 (100%)	0 (0%)	1 (100%)	1 (0%)
Cruising without rotation	1	1 (100%)	0 (0%)	1 (100%)	0 (0%)	1 (100%)	1 (0%)
Controlled lowering through standing	0	-	-	-	-	-	-
Cruising with rotation	0	-	-	-	-	-	-
Stands alone	0	-	-	-	-	-	-
Early stepping	0	-	-	-	-	-	-
Standing from modified squat	0	-	-	-	-	-	-
Standing from quadruped position	0	-	-	-	-	-	-
Walks alone	0	-	-	-	-	-	-
Squat	0	-	-	-	-	-	-
38 Movements	**78**	**59 (76%)**	**19 (24%)**	**58 (74%)**	**20 (26%)**	**53 (68%)**	**25 (32%)**

"Baby Go" Version 2.0, updated on 2024-10-04 with the addition of assessment, follow-up, and education features and reminders, and a video editor

doi:10.6342/NTU202504540

Table 6. Agreement of movement-based assessment results for application version 2.1 (Cut-off 2025-02-18)

Movements	No. uploaded	Parent vs. PT		AI vs. PT		Parent vs. AI	
		Agreed	Disagreed	Agreed	Disagreed	Agreed	Disagreed
Forearm support (1)	23	15 (65%)	8 (35%)	14 (61%)	9 (39%)	9 (39%)	14 (61%)
Forearm support (2)	14	12 (86%)	2 (14%)	12 (86%)	2 (14%)	13 (93%)	1 (7%)
Extend arm support	14	8 (57%)	6 (43%)	8 (57%)	6 (43%)	5 (36%)	9 (64%)
Reach from forearm support	4	3 (75%)	1 (25%)	3 (75%)	1 (25%)	3 (75%)	1 (25%)
Pivoting	2	1 (50%)	1 (50%)	2 (100%)	0 (0%)	1 (50%)	1 (50%)
Four-point kneeling (1)	2	2 (100%)	0 (0%)	2 (100%)	0 (0%)	2 (100%)	0 (0%)
Reciprocal crawling	5	2 (40%)	3 (60%)	3 (60%)	2 (40%)	3 (60%)	2 (40%)
Four-point kneeling to sitting	1	1 (100%)	0 (0%)	1 (100%)	0 (0%)	1 (100%)	0 (0%)
Reciprocal creeping (1)	1	1 (100%)	0 (0%)	0 (0%)	1 (100%)	0 (0%)	1 (100%)
Reaching from extended arm support	2	2 (100%)	0 (0%)	0 (0%)	2 (100%)	0 (0%)	2 (100%)
Reciprocal creeping (2)	1	1 (100%)	0 (0%)	1 (100%)	0 (0%)	1 (100%)	0 (0%)
Supine lying (3)	15	15 (100%)	0 (0%)	14 (93%)	1 (7%)	14 (93%)	1 (7%)
Supine lying (4)	15	15 (100%)	0 (0%)	10 (67%)	5 (33%)	10 (67%)	5 (33%)
Hands to knees	7	4 (57%)	3 (43%)	5 (71%)	2 (29%)	7 (100%)	0 (0%)
Hands to feet	7	6 (86%)	1 (14%)	6 (86%)	1 (14%)	6 (86%)	1 (14%)
Rolling supine to prone with rotation	2	2 (100%)	0 (0%)	2 (100%)	0 (0%)	2 (100%)	0 (0%)
Sitting with support	23	23 (100%)	0 (0%)	6 (26%)	17 (74%)	6 (26%)	17 (74%)
Sitting with propped arms	14	8 (57%)	6 (43%)	10 (71%)	4 (29%)	7 (50%)	7 (50%)
Pull to sit	9	9 (100%)	0 (0%)	8 (89%)	1 (11%)	8 (89%)	1 (11%)
Sitting with arm support	9	8 (89%)	1 (11%)	8 (89%)	1 (11%)	8 (89%)	1 (11%)
Sitting without arm support (1)	8	7 (88%)	1 (13%)	5 (63%)	3 (38%)	4 (50%)	4 (50%)
Reach with rotation in sitting	1	1 (100%)	0 (0%)	0 (0%)	1 (100%)	0 (0%)	1 (100%)
Sitting to four-point kneeling	2	2 (100%)	0 (0%)	1 (50%)	1 (50%)	1 (50%)	1 (50%)
Sitting without arm support (2)	1	0 (0%)	1 (100%)	0 (0%)	1 (100%)	0 (0%)	1 (100%)
Supported standing (2)	9	9 (100%)	0 (0%)	7 (78%)	2 (22%)	7 (78%)	2 (22%)
Supported standing (3)	0	-	-	-	-	-	-
Pull to stand with support	3	3 (100%)	0 (0%)	0 (0%)	3 (100%)	0 (0%)	3 (100%)
Pulls to stand/stands	0	-	-	-	-	-	-
Supported standing with rotation	1	1 (100%)	0 (0%)	0 (0%)	1 (100%)	0 (0%)	1 (100%)
Cruising without rotation	0	-	-	-	-	-	-
Controlled lowering through standing	2	1 (50%)	1 (50%)	1 (50%)	1 (50%)	1 (50%)	1 (50%)
Cruising with rotation	0	-	-	-	-	-	-
Stands alone	2	2 (100%)	0 (0%)	1 (50%)	1 (50%)	1 (50%)	1 (50%)
Early stepping	0	-	-	-	-	-	-
Standing from modified squat	0	-	-	-	-	-	-
Standing from quadruped position	0	-	-	-	-	-	-
Walks alone	0	-	-	-	-	-	-
Squat	0	-	-	-	-	-	-
38 Movements	**199**	**164 (82%)**	**35 (18%)**	**130 (65%)**	**69 (35%)**	**120 (60%)**	**79 (40%)**

"Baby Go" Version 2.1, updated on 2024-12-05 with changes in uploading interface and wording revision for recording time duration

doi:10.6342/NTU202504540

Table 7. Agreement of movement-based assessment results for application version 2.2 (Cut-off 2025-03-08)

Movements	No. uploaded	Parent vs. PT		AI vs. PT		Parent vs. AI	
		Agreed	Disagreed	Agreed	Disagreed	Agreed	Disagreed
Forearm support (1)	6	4 (67%)	2 (33%)	5 (83%)	1 (17%)	3 (50%)	3 (50%)
Forearm support (2)	3	3 (100%)	0 (0%)	3 (100%)	0 (0%)	3 (100%)	0 (0%)
Extend arm support	2	2 (100%)	0 (0%)	2 (100%)	0 (0%)	2 (100%)	0 (0%)
Reach from forearm support	3	2 (67%)	1 (33%)	1 (33%)	2 (67%)	1 (33%)	2 (67%)
Pivoting	1	1 (100%)	0 (0%)	1 (100%)	0 (0%)	1 (100%)	0 (0%)
Four-point kneeling (1)	-	-	-	-	-	-	-
Reciprocal crawling	1	0 (0%)	1 (100%)	0 (0%)	1 (100%)	0 (0%)	1 (100%)
Four-point kneeling to sitting	0	-	-	-	-	-	-
Reciprocal creeping (1)	0	-	-	-	-	-	-
Reaching from extended arm support	0	-	-	-	-	-	-
Reciprocal creeping (2)	0	-	-	-	-	-	-
Supine lying (3)	2	2 (100%)	0 (0%)	2 (100%)	0 (0%)	2 (100%)	0 (0%)
Supine lying (4)	4	4 (100%)	0 (0%)	4 (100%)	0 (0%)	4 (100%)	0 (0%)
Hands to knees	2	2 (100%)	0 (0%)	2 (100%)	0 (0%)	2 (100%)	0 (0%)
Hands to feet	1	1 (100%)	0 (0%)	1 (100%)	0 (0%)	1 (100%)	0 (0%)
Rolling supine to prone with rotation	1	1 (100%)	0 (0%)	1 (100%)	0 (0%)	1 (100%)	0 (0%)
Sitting with support	6	6 (100%)	0 (0%)	5 (83%)	1 (17%)	5 (83%)	1 (17%)
Sitting with propped arms	1	0 (0%)	1 (100%)	1 (100%)	0 (0%)	0 (0%)	1 (100%)
Pull to sit	0	-	-	-	-	-	-
Sitting with arm support	0	-	-	-	-	-	-
Sitting without arm support (1)	1	1 (100%)	0 (0%)	1 (100%)	0 (0%)	1 (100%)	0 (0%)
Reach with rotation in sitting	0	-	-	-	-	-	-
Sitting to four-point kneeling	0	-	-	-	-	-	-
Sitting without arm support (2)	1	0 (0%)	1 (100%)	1 (100%)	0 (0%)	0 (0%)	1 (100%)
Supported standing (2)	2	1 (50%)	1 (50%)	1 (50%)	1 (50%)	2 (100%)	0 (0%)
Supported standing (3)	1	0 (0%)	1 (100%)	1 (100%)	0 (0%)	0 (0%)	1 (100%)
Pull to stand with support	0	-	-	-	-	-	-
Pulls to stand/stands	0	-	-	-	-	-	-
Supported standing with rotation	0	-	-	-	-	-	-
Cruising without rotation	0	-	-	-	-	-	-
Controlled lowering through standing	0	-	-	-	-	-	-
Cruising with rotation	1	1 (100%)	0 (0%)	1 (100%)	0 (0%)	1 1 (100%)	0 (0%)
Stands alone	0	-	-	-	-	-	-
Early stepping	0	-	-	-	-	-	-
Standing from modified squat	0	-	-	-	-	-	-
Standing from quadruped position	0	-	-	-	-	-	-
Walks alone	0	-	-	-	-	-	-
Squat	0	-	-	-	-	-	-
38 Movements	**39**	**31 (76%)**	**11 (21%)**	**33 (85%)**	**6 (15%)**	**29 (74%)**	**10 (26%)**

"Baby Go" Version 2.2, updated on 2025-02-18 with changes in the logic for advanced movements and added guideline to identify confusing movements.

doi:10.6342/NTU202504540

Table 8. Agreement of movement-based assessment results for application version 2.3 (Cut-off 2025-06-30)

Movements	No. uploaded	Parent vs. PT		AI vs. PT		Parent vs. AI	
		Agreed	**Disagreed**	**Agreed**	**Disagreed**	**Agreed**	**Disagreed**
Forearm support (1)	18	5 (28%)	13 (72%)	13 (72%)	5 (28%)	7 (39%)	11 (61%)
Forearm support (2)	22	17 (77%)	5 (23%)	15 (68%)	7 (32%)	15 (68%)	7 (32%)
Extend arm support	27	12 (44%)	15 (56%)	25 (93%)	2 (7%)	16 (59%)	11 (41%)
Reach from forearm support	21	11 (52%)	10 (48%)	9 (43%)	12 (57%)	8 (38%)	13 (62%)
Pivoting	15	12 (80%)	3 (20%)	13 (87%)	2 (13%)	11 (73%)	4 (27%)
Four-point kneeling (1)	23	15 (65%)	8 (35%)	12 (52%)	11 (48%)	12 (52%)	11 (48%)
Reciprocal crawling	14	6 (43%)	8 (57%)	9 (64%)	5 (36%)	8 (57%)	6 (43%)
Four-point kneeling to sitting	22	20 (91%)	2 (9%)	12 (55%)	10 (45%)	12 (55%)	10 (45%)
Reciprocal creeping (1)	15	12 (80%)	3 (20%)	12 (80%)	3 (20%)	13 (87%)	2 (13%)
Reaching from extended arm support	21	16 (76%)	5 (24%)	4 (19%)	17 (81%)	2 (10%)	19 (90%)
Reciprocal creeping (2)	16	14 (88%)	2 (13%)	11 (69%)	5 (31%)	11 (69%)	5 (31%)
Supine lying (3)	9	6 (67%)	3 (33%)	9 (100%)	0 (0%)	6 (67%)	3 (33%)
Supine lying (4)	16	15 (94%)	1 (6%)	14 (88%)	2 (13%)	13 (81%)	3 (19%)
Hands to knees	16	15 (94%	1 (6%)	14 (88%)	2 (13%)	14 (88%)	2 (13%)
Hands to feet	21	21 (100%	0 (0%)	21 (100%)	0 (0%)	21 (100%)	0 (0%)
Rolling supine to prone with rotation	21	20 (95%)	1 (5%)	16 (76%)	5 (24%)	17 (81%)	4 (19%)
Sitting with support	14	14 (100%	0 (0%)	8 (57%)	6 (43%)	8 (57%)	6 (43%)
Sitting with propped arms	24	14 (58%)	10 (42%)	21 (88%)	3 (13%)	19 (79%)	5 (21%)
Pull to sit	20	19 (95%)	1 (5%)	16 (80%)	4 (20%)	16 (80%)	4 (20%)
Sitting with arm support	28	24 (86%)	4 (14%)	21 (75%)	7 (25%)	21 (75%)	7 (25%)
Sitting without arm support (1)	30	25 (83%)	5 (17%)	19 (63%)	11 (37%)	18 (60%)	12 (40%)
Reach with rotation in sitting	31	16 (52%)	15 (48%)	14 (45%)	17 (55%)	9 (29%)	22 (71%)
Sitting to four-point kneeling	23	18 (78%)	5 (22%)	11 (48%)	12 (52%)	14 (61%)	9 (39%)
Sitting without arm support (2)	20	8 (40%)	12 (60%)	12 (60%)	8 (40%)	12 (60%)	8 (40%)
Supported standing (2)	18	18 (100%	0 (0%)	15 (83%)	3 (17%)	15 (83%)	3 (17%)
Supported standing (3)	11	9 (82%)	2 (18%)	1 (9%)	10 (91%)	2 (18%)	9 (82%)
Pull to stand with support	22	15 (68%)	7 (32%)	11 (50%)	11 (50%)	11 (50%)	11 (50%)
Pulls to stand/stands	26	19 (73%)	7 (27%)	11 (42%)	15 (58%)	11 (42%)	15 (58%)
Supported standing with rotation	14	11 (79%)	3 (21%)	6 (43%)	8 (57%)	5 (36%)	9 (64%)
Cruising without rotation	11	8 (73%)	3 (27%)	7 (64%)	4 (36%)	8 (73%)	3 (27%)
Controlled lowering through standing	9	8 (89%)	1 (11%)	6 (67%)	3 (33%)	6 (67%)	3 (33%)
Cruising with rotation	3	1 (33%)	2 (67%)	1 (33%)	2 (67%)	0 (0%)	3 (100%)
Stands alone	5	5 (100%)	0 (0%)	3 (60%)	2 (40%)	3 (60%)	2 (40%)
Early stepping	4	4 (100%)	0 (0%)	3 (75%)	1 (25%)	3 (75%)	1 (25%)
Standing from modified squat	3	3 (100%)	0 (0%)	1 (33%)	2 (67%)	1 (33%)	2 (67%)
Standing from quadruped position	3	3 (100%)	0 (0%)	3 (100%)	0 (0%)	3 (100%)	0 (0%)
Walks alone	2	2 (100%)	0 (0%)	2 (100%)	0 (0%)	2 (100%)	0 (0%)
Squat	2	2 (100%)	0 (0%)	2 (100%)	0 (0%)	2 (100%)	0 (0%)
38 Movements	**620**	**463 (75%)**	**157 (25%)**	**403 (65%)**	**217 (35%)**	**375 (60%)**	**245 (40%)**

"Baby Go" version 2.3, updated on 2025-03-08 with accommodation for the vertical position problem during uploading videos

doi:10.6342/NTU202504540

Table 9. Agreement of movement-based assessment results for application version 2.0 - 2.3
(All data, with portrait and incorrect orientation included, Cut-off 2025-06-30)

Movements	No. uploaded	Parent vs. PT		AI vs. PT		Parent vs. AI	
		Agreed	Disagreed	Agreed	Disagreed	Agreed	Disagreed
Forearm support (1)	63	32 (51%)	31 (49%)	43 (68%)	20 (32%)	27 (43%)	36 (57%)
Forearm support (2)	40	33 (83%)	7 (18%)	31 (78%)	9 (23%)	32 (80%)	8 (20%)
Extend arm support	51	26 (51%)	25 (49%)	40 (78%)	11 (22%)	25 (49%)	26 (51%)
Reach from forearm support	29	17 (59%)	12 (41%)	14 (48%)	15 (52%)	13 (45%)	16 (55%)
Pivoting	19	15 (79%)	4 (21%)	17 (89%)	2 (11%)	14 (74%)	5 (26%)
Four-point kneeling (1)	26	18 (69%)	8 (31%)	14 (54%)	12 (46%)	14 (54%)	12 (46%)
Reciprocal crawling	21	9 (43%)	12 (57%)	13 (62%)	8 (38%)	12 (57%)	9 (43%)
Four-point kneeling to sitting	24	22 (92%)	2 (8%)	14 (58%)	10 (42%)	14 (58%)	10 (42%)
Reciprocal creeping (1)	17	14 (82%)	3 (18%)	13 (76%)	4 (24%)	14 (82%)	3 (18%)
Reaching from extended arm support	24	19 (79%)	5 (21%)	4 (17%)	20 (83%)	2 (8%)	22 (92%)
Reciprocal creeping (2)	18	16 (89%)	2 (11%)	13 (72%)	5 (28%)	13 (72%)	5 (28%)
Supine lying (3)	36	32 (89%)	4 (11%)	33 (92%)	3 (8%)	30 (83%)	6 (17%)
Supine lying (4)	41	37 (90%)	4 (10%)	33 (80%)	8 (20%)	32 (78%)	9 (22%)
Hands to knees	27	23 (85%)	4 (15%)	23 (85%)	4 (15%)	25 (93%)	2 (7%)
Hands to feet	33	32 (97%)	1 (3%)	32 (97%)	1 (3%)	32 (97%)	1 (3%)
Rolling supine to prone with rotation	26	25 (96%)	1 (4%)	21 (81%)	5 (19%)	22 (85%)	4 (15%)
Sitting with support	52	51 (98%)	1 (2%)	23 (44%)	29 (56%)	24 (46%)	28 (54%)
Sitting with propped arms	41	23 (56%)	18 (44%)	34 (83%)	7 (17%)	27 (66%)	14 (34%)
Pull to sit	30	29 (97%)	1 (3%)	24 (80%)	6 (20%)	24 (80%)	6 (20%)
Sitting with arm support	38	33 (87%)	5 (13%)	30 (79%)	8 (21%)	30 (79%)	8 (21%)
Sitting without arm support (1)	40	33 (83%)	7 (18%)	25 (63%)	15 (38%)	24 (60%)	16 (40%)
Reach with rotation in sitting	32	17 (53%)	15 (47%)	14 (44%)	18 (56%)	9 (28%)	23 (72%)
Sitting to four-point kneeling	26	21 (81%)	5 (19%)	13 (50%)	13 (50%)	16 (62%)	10 (38%)
Sitting without arm support (2)	23	9 (39%)	14 (61%)	14 (61%)	9 (39%)	13 (57%)	10 (43%)
Supported standing (2)	30	29 (97%)	1 (3%)	24 (80%)	6 (20%)	25 (83%)	5 (17%)
Supported standing (3)	13	10 (77%)	3 (23%)	3 (23%)	10 (77%)	3 (23%)	10 (77%)
Pull to stand with support	25	18 (72%)	7 (28%)	11 (44%)	14 (56%)	11 (44%)	14 (56%)
Pulls to stand/stands	27	20 (74%)	7 (26%)	12 (44%)	15 (56%)	12 (44%)	15 (56%)
Supported standing with rotation	16	13 (81%)	3 (19%)	7 (44%)	9 (56%)	6 (38%)	10 (63%)
Cruising without rotation	12	9 (75%)	3 (25%)	8 (67%)	4 (33%)	9 (75%)	3 (25%)
Controlled lowering through standing	11	9 (82%)	2 (18%)	7 (64%)	4 (36%)	7 (64%)	4 (36%)
Cruising with rotation	4	2 (50%)	2 (50%)	2 (50%)	2 (50%)	1 (25%)	3 (75%)
Stands alone	7	7 (100%)	0 (0%)	4 (57%)	3 (43%)	4 (57%)	3 (43%)
Early stepping	4	4 (100%)	0 (0%)	3 (75%)	1 (25%)	3 (75%)	1 (25%)
Standing from modified squat	3	3 (100%)	0 (0%)	1 (33%)	2 (67%)	1 (33%)	2 (67%)
Standing from quadruped position	3	3 (100%)	0 (0%)	3 (100%)	0 (0%)	3 (100%)	0 (0%)
Walks alone	2	2 (100%)	0 (0%)	2 (100%)	0 (0%)	2 (100%)	0 (0%)
Squat	2	2 (100%)	0 (0%)	2 (100%)	0 (0%)	2 (100%)	0 (0%)
38 Movements	**936**	**717 (77%)**	**219 (23%)**	**624 (67%)**	**312 (33%)**	**577 (62%)**	**359 (38%)**

doi:10.6342/NTU202504540

Table 10. Agreement of movement-based assessment results based on AI model I (All data before 2025-01-14)

Movements	No. uploaded	Parent vs. PT		AI vs. PT		Parent vs. AI	
		Agreed	Disagreed	Agreed	Disagreed	Agreed	Disagreed
Forearm support (1)	31	18 (58%)	13 (42%)	20 (65%)	11 (35%)	15 (48%)	16 (52%)
Forearm support (2)	10	8 (80%)	2 (20%)	9 (90%)	1 (10%)	10 (100%)	0 (0%)
Extend arm support	14	6 (43%)	8 (57%)	10 (71%)	4 (29%)	4 (29%)	10 (71%)
Reach from forearm support	2	2 (100%)	0 (0%)	2 (100%)	0 (0%)	2 (100%)	0 (0%)
Pivoting	1	1 (100%)	0 (0%)	1 (100%)	0 (0%)	1 (100%)	0 (0%)
Four-point kneeling (1)	2	2 (100%)	0 (0%)	1 (50%)	1 (50%)	1 (50%)	1 (50%)
Reciprocal crawling	3	1 (33%)	2 (67%)	2 (67%)	1 (33%)	1 (33%)	2 (67%)
Four-point kneeling to sitting	1	1 (100%)	0 (0%)	1 (100%)	0 (0%)	1 (100%)	0 (0%)
Reciprocal creeping (1)	1	1 (100%)	0 (0%)	1 (100%)	0 (0%)	1 (100%)	0 (0%)
Reaching from extended arm support	3	3 (100%)	0 (0%)	0 (0%)	3 (100%)	0 (0%)	3 (100%)
Reciprocal creeping (2)	1	1 (100%)	0 (0%)	1 (100%)	0 (0%)	1 (100%)	0 (0%)
Supine lying (3)	21	20 (95%)	1 (5%)	18 (86%)	3 (14%)	18 (86%)	3 (14%)
Supine lying (4)	13	10 (77%)	3 (23%)	9 (69%)	4 (31%)	9 (69%)	4 (31%)
Hands to knees	8	5 (63%)	3 (38%)	6 (75%)	2 (25%)	8 (100%)	0 (0%)
Hands to feet	7	7 (100%)	0 (0%)	6 (86%)	1 (14%)	6 (86%)	1 (14%)
Rolling supine to prone with rotation	2	2 (100%)	0 (0%)	2 (100%)	0 (0%)	2 (100%)	0 (0%)
Sitting with support	24	23 (96%)	1 (4%)	9 (38%)	15 (63%)	10 (42%)	14 (58%)
Sitting with propped arms	8	6 (75%)	2 (25%)	8 (100%)	0 (0%)	6 (75%)	2 (25%)
Pull to sit	7	7 (100%)	0 (0%)	6 (86%)	1 (14%)	6 (86%)	1 (14%)
Sitting with arm support	6	6 (100%)	0 (0%)	5 (83%)	1 (17%)	5 (83%)	1 (17%)
Sitting without arm support (1)	8	6 (75%)	2 (25%)	4 (50%)	4 (50%)	4 (50%)	4 (50%)
Reach with rotation in sitting	0	-	-	-	-	-	-
Sitting to four-point kneeling	2	2 (100%)	0 (0%)	2 (100%)	0 (0%)	2 (100%)	0 (0%)
Sitting without arm support (2)	2	1 (50%)	1 (50%)	1 (50%)	1 (50%)	1 (50%)	1 (50%)
Supported standing (2)	8	8 (100%)	0 (0%)	6 (75%)	2 (25%)	6 (75%)	2 (25%)
Supported standing (3)	1	1 (100%)	0 (0%)	1 (100%)	0 (0%)	1 (100%)	0 (0%)
Pull to stand with support	1	1 (100%)	0 (0%)	0 (0%)	1 (100%)	0 (0%)	1 (100%)
Pulls to stand/stands	1	1 (100%)	0 (0%)	1 (100%)	0 (0%)	1 (100%)	0 (0%)
Supported standing with rotation	2	2 (100%)	0 (0%)	1 (50%)	1 (50%)	1 (50%)	1 (50%)
Cruising without rotation	1	1 (100%)	0 (0%)	1 (100%)	0 (0%)	1 (100%)	0 (0%)
Controlled lowering through standing	2	1 (50%)	1 (50%)	1 (50%)	1 (50%)	1 (50%)	1 (50%)
Cruising with rotation	0	-	-	-	-	-	-
Stands alone	1	1 (100%)	0 (0%)	0 (0%)	1 (100%)	0 (0%)	1 (100%)
Early stepping	0	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
Standing from modified squat	0	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
Standing from quadruped position	0	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
Walks alone	0	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
Squat	0	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
38 Movements	**194**	**155 (80%)**	**39 (20%)**	**135 (70%)**	**59 (30%)**	**125 (64%)**	**69 (36%)**

Before 2025-01-14, the application applied the AI model trained on 1,396 data samples from home videos.

doi:10.6342/NTU202504540

Table 11. Agreement of movement-based assessment results based on AI model II (All data after (All data after 2025-01-14, Cut-off 2025-06-30)

Movements	No. uploaded	Parent vs. PT		AI vs. PT		Parent vs. AI	
		Agreed	Disagreed	Agreed	Disagreed	Agreed	Disagreed
Forearm support (1)	32	14 (44%)	18 (60%)	23 (72%)	9 (28%)	12 (38%)	20 (63%)
Forearm support (2)	30	25 (83%)	5 (14%)	22 (73%)	8 (27%)	22 (73%)	8 (27%)
Extend arm support	37	20 (54%)	17 (63%)	30 (81%)	7 (19%)	21 (57%)	16 (43%)
Reach from forearm support	27	15 (56%)	12 (67%)	12 (44%)	15 (56%)	11 (41%)	16 (59%)
Pivoting	18	14 (78%)	4 (17%)	16 (89%)	2 (11%)	13 (72%)	5 (28%)
Four-point kneeling (1)	24	16 (67%)	8 (44%)	13 (54%)	11 (46%)	13 (54%)	11 (46%)
Reciprocal crawling	18	8 (44%)	10 (43%)	11 (61%)	7 (39%)	11 (61%)	7 (39%)
Four-point kneeling to sitting	23	21 (91%)	2 (13%)	13 (57%)	10 (43%)	13 (57%)	10 (43%)
Reciprocal creeping (1)	16	13 (81%)	3 (14%)	12 (75%)	4 (25%)	13 (81%)	3 (19%)
Reaching from extended arm support	21	16 (76%)	5 (29%)	4 (19%)	17 (81%)	2 (10%)	19 (90%)
Reciprocal creeping (2)	17	15 (88%)	2 (13%)	12 (71%)	5 (29%)	12 (71%)	5 (29%)
Supine lying (3)	15	12 (80%)	3 (11%)	15 (100%)	0 (0%)	12 (80%)	3 (20%)
Supine lying (4)	28	27 (96%)	1 (5%)	24 (86%)	4 (14%)	23 (82%)	5 (18%)
Hands to knees	19	18 (95%)	1 (4%)	17 (89%)	2 (11%)	17 (89%)	2 (11%)
Hands to feet	26	25 (96%)	1 (4%)	26 (100%)	0 (0%)	26 (100%)	0 (0%)
Rolling supine to prone with rotation	24	23 (96%)	1 (4%)	19 (79%)	5 (21%)	20 (83%)	4 (17%)
Sitting with support	28	28 (100%)	0 (0%)	14 (50%)	14 (50%)	14 (50%)	14 (50%)
Sitting with propped arms	33	17 (52%)	16 (70%)	26 (79%)	7 (21%)	21 (64%)	12 (36%)
Pull to sit	23	22 (96%)	1 (3%)	18 (78%)	5 (22%)	18 (78%)	5 (22%)
Sitting with arm support	32	27 (84%)	5 (16%)	25 (78%)	7 (22%)	25 (78%)	7 (22%)
Sitting without arm support (1)	32	27 (84%)	5 (16%)	21 (66%)	11 (34%)	20 (63%)	12 (38%)
Reach with rotation in sitting	32	17 (53%)	15 (63%)	14 (44%)	18 (56%)	9 (28%)	23 (72%)
Sitting to four-point kneeling	24	19 (79%)	5 (24%)	11 (46%)	13 (54%)	14 (58%)	10 (42%)
Sitting without arm support (2)	21	8 (38%)	13 (59%)	13 (62%)	8 (38%)	12 (57%)	9 (43%)
Supported standing (2)	22	21 (95%)	1 (8%)	18 (82%)	4 (18%)	19 (86%)	3 (14%)
Supported standing (3)	12	9 (75%)	3 (13%)	2 (17%)	10 (83%)	2 (17%)	10 (83%)
Pull to stand with support	24	17 (71%)	7 (27%)	11 (46%)	13 (54%)	11 (46%)	13 (54%)
Pulls to stand/stands	26	19 (73%)	7 (50%)	11 (42%)	15 (58%)	11 (42%)	15 (58%)
Supported standing with rotation	14	11 (79%)	3 (27%)	6 (43%)	8 (57%)	5 (36%)	9 (64%)
Cruising without rotation	11	8 (73%)	3 (33%)	7 (64%)	4 (36%)	8 (73%)	3 (27%)
Controlled lowering through standing	9	8 (89%)	1 (25%)	6 (67%)	3 (33%)	6 (67%)	3 (33%)
Cruising with rotation	4	2 (50%)	2 (33%)	2 (50%)	2 (50%)	1 (25%)	3 (75%)
Stands alone	6	6 (100%)	0 (0%)	4 (67%)	2 (33%)	4 (67%)	2 (33%)
Early stepping	4	4 (100%)	0 (0%)	3 (75%)	1 (25%)	3 (75%)	1 (25%)
Standing from modified squat	3	3 (100%)	0 (0%)	1 (33%)	2 (67%)	1 (33%)	2 (67%)
Standing from quadruped position	3	3 (100%)	0 (0%)	3 (100%)	0 (0%)	3 (100%)	0 (0%)
Walks alone	2	2 (100%)	0 (0%)	2 (100%)	0 (0%)	2 (100%)	0 (0%)
Squat	2	2 (100%)	0 (0%)	2 (100%)	0 (0%)	2 (100%)	0 (0%)
38 Movements	**742**	**562 (76%)**	**180 (24%)**	**489 (66%)**	**253 (34%)**	**452 (61%)**	**290 (39%)**

After 2025-01-14, the application applied the AI model trained on 474 videos with 7,234 data samples from laboratory and 461 videos with 1,419 data samples from home.

doi:10.6342/NTU202504540

Table 12. Agreement of movement-based assessment results for application version 2.0 to 2.3 using videos with portrait mode only, (Cut-off 2025-06-30)

Movements	No. uploaded	Parent vs. PT		AI vs. PT		Parent vs. AI	
		Agreed	Disagreed	Agreed	Disagreed	Agreed	Disagreed
Forearm support (1)	9	1 (100%)	0 (0%)	1 (100%)	0 (0%)	1 (100%)	0 (0%)
Forearm support (2)	3	4 (44%)	5 (56%)	7 (78%)	2 (22%)	3 (33%)	6 (67%)
Extend arm support	7	3 (50%)	3 (50%)	4 (67%)	2 (33%)	3 (50%)	3 (50%)
Reach from forearm support	8	2 (40%)	3 (60%)	2 (40%)	3 (60%)	5 (100%)	0 (0%)
Pivoting	6	3 (75%)	1 (25%)	3 (75%)	1 (25%)	4 (100%)	0 (0%)
Four-point kneeling (1)	10	5 (38%)	8 (100%)	3 (38%)	5 (63%)	5 (63%)	3 (38%)
Reciprocal crawling	5	2 (67%)	1 (33%)	3 (100%)	0 (0%)	2 (67%)	1 (33%)
Four-point kneeling to sitting	11	2 (20%)	8 (80%)	2 (20%)	8 (80%)	10 (100%)	0 (0%)
Reciprocal creeping (1)	4	5 (38%)	8 (62%)	5 (38%)	8 (62%)	3 (23%)	10 (77%)
Reaching from extended arm support	6	1 (11%)	8 (89%)	1 (11%)	8 (89%)	5 (56%)	4 (44%)
Reciprocal creeping (2)	4	2 (67%)	1 (33%)	2 (67%)	1 (33%)	3 (100%)	0 (0%)
Supine lying (3)	12	13 (100%)	0 (0%)	13 (100%)	0 (0%)	12 (92%)	1 (8%)
Supine lying (4)	11	7 (70%)	3 (30%)	7 (70%)	3 (30%)	10 (100%)	0 (0%)
Hands to knees	9	5 (83%)	1 (17%)	5 (83%)	1 (17%)	6 (100%)	0 (0%)
Hands to feet	9	10 (100%)	0 (0%)	10 (100%)	0 (0%)	9 (90%)	1 (10%)
Rolling supine to prone with rotation	9	4 (50%)	4 (50%)	4 (50%)	4 (50%)	8 (100%)	0 (0%)
Sitting with support	15	7 (47%)	8 (53%)	7 (47%)	8 (53%)	15 (100%)	0 (0%)
Sitting with propped arms	13	5 (100%)	0 (0%)	5 (100%)	0 (0%)	5 (100%)	0 (0%)
Pull to sit	5	4 (80%)	1 (20%)	4 (80%)	1 (20%)	5 (100%)	0 (0%)
Sitting with arm support	16	11 (50%)	11 (50%)	17 (77%)	5 (23%)	12 (55%)	10 (45%)
Sitting without arm support (1)	21	12 (48%)	13 (52%)	13 (52%)	12 (48%)	14 (56%)	11 (44%)
Reach with rotation in sitting	12	2 (40%)	3 (60%)	2 (40%)	3 (60%)	5 (100%)	0 (0%)
Sitting to four-point kneeling	10	3 (23%)	10 (77%)	4 (31%)	9 (69%)	8 (62%)	5 (38%)
Sitting without arm support (2)	11	3 (75%)	1 (25%)	4 (100%)	0 (0%)	2 (50%)	2 (50%)
Supported standing (2)	13	10 (67%)	5 (33%)	10 (67%)	5 (33%)	13 (87%)	2 (13%)
Supported standing (3)	8	2 (40%)	3 (60%)	2 (40%)	3 (60%)	5 (100%)	0 (0%)
Pull to stand with support	7	2 (50%)	2 (50%)	2 (50%)	2 (50%)	3 (75%)	1 (25%)
Pulls to stand/stands	15	4 (33%)	8 (67%)	4 (33%)	8 (67%)	10 (83%)	2 (17%)
Supported standing with rotation	9	5 (71%)	2 (29%)	5 (71%)	2 (29%)	7 (100%)	0 (0%)
Cruising without rotation	6	2 (67%)	1 (33%)	2 (67%)	1 (33%)	3 (100%)	0 (0%)
Controlled lowering through standing	4	2 (100%)	0 (0%)	2 (100%)	0 (0%)	2 (100%)	0 (0%)
Cruising with rotation	3	1 (50%)	1 (50%)	1 (50%)	1 (50%)	**2 (100%)**	0 (0%)
Stands alone	4	2 (50%)	2 (50%)	2 (50%)	2 (50%)	4 (100%)	0 (0%)
Early stepping	3	2 (67%)	1 (33%)	2 (67%)	1 (33%)	3 (100%)	0 (0%)
Standing from modified squat	2	0 (0%)	2 (100%)	0 (0%)	2 (100%)	2 (100%)	0 (0%)
Standing from quadruped position	1	1 (100%)	0 (0%)	1 (100%)	0 (0%)	1 (100%)	0 (0%)
Walks alone	1	1 (100%)	0 (0%)	1 (100%)	0 (0%)	1 (100%)	0 (0%)
Squat	1	1 (100%)	0 (0%)	1 (100%)	0 (0%)	1 (100%)	0 (0%)
38 Movements	**303**	**146 (53%)**	**128 (47%)**	**163 (59%)**	**111 (41%)**	**212 (77%)**	**62 (23%)**

doi:10.6342/NTU202504540

Table 13. Agreement of movement-based assessment results for application version 2.0 to 2.3 using videos with incorrect orientation only (Cut-off 2025-06-30)

Movements	No. uploaded	Parent vs. PT		AI vs. PT		Parent vs. AI	
		Agreed	Disagreed	Agreed	Disagreed	Agreed	Disagreed
Forearm support (1)	20	9 (45%)	11 (55%)	13 (65%)	7 (35%)	8 (40%)	12 (60%)
Forearm support (2)	5	5 (100%)	0 (0%)	4 (80%)	1 (20%)	4 (80%)	1 (20%)
Extend arm support	7	5 (71%)	2 (29%)	2 (29%)	5 (71%)	0 (0%)	7 (100%)
Reach from forearm support	1	1 (100%)	0 (0%)	1 (100%)	0 (0%)	1 (100%)	0 (0%)
Pivoting	2	2 (100%)	0 (0%)	2 (100%)	0 (0%)	2 (100%)	0 (0%)
Four-point kneeling (1)	3	3 (100%)	0 (0%)	1 (33%)	2 (67%)	1 (33%)	2 (67%)
Reciprocal crawling	5	2 (40%)	3 (60%)	0 (0%)	5 (100%)	1 (20%)	4 (80%)
Four-point kneeling to sitting	1	1 (100%)	0 (0%)	1 (100%)	0 (0%)	1 (100%)	0 (0%)
Reciprocal creeping (1)	2	2 (100%)	0 (0%)	1 (50%)	1 (50%)	1 (50%)	1 (50%)
Reaching from extended arm support	4	4 (100%)	0 (0%)	0 (0%)	4 (100%)	0 (0%)	4 (100%)
Reciprocal creeping (2)	2	1 (50%)	1 (50%)	0 (0%)	2 (100%)	0 (0%)	2 (100%)
Supine lying (3)	8	8 (100%)	0 (0%)	6 (75%)	2 (25%)	6 (75%)	2 (25%)
Supine lying (4)	8	8 (100%)	0 (0%)	5 (63%)	3 (38%)	5 (63%)	3 (38%)
Hands to knees	2	2 (100%)	0 (0%)	2 (100%)	0 (0%)	2 (100%)	0 (0%)
Hands to feet	2	2 (100%)	0 (0%)	2 (100%)	0 (0%)	2 (100%)	0 (0%)
Rolling supine to prone with rotation	3	3 (100%)	0 (0%)	3 (100%)	0 (0%)	3 (100%)	0 (0%)
Sitting with support	15	14 (93%)	1 (7%)	3 (20%)	12 (80%)	4 (27%)	11 (73%)
Sitting with propped arms	9	5 (56%)	4 (44%)	4 (44%)	5 (56%)	2 (22%)	7 (78%)
Pull to sit	4	4 (100%)	0 (0%)	3 (75%)	1 (25%)	3 (75%)	1 (25%)
Sitting with arm support	1	1 (100%)	0 (0%)	0 (0%)	1 (100%)	0 (0%)	1 (100%)
Sitting without arm support (1)	8	8 (100%)	0 (0%)	3 (38%)	5 (63%)	3 (38%)	5 (63%)
Reach with rotation in sitting	3	1 (33%)	2 (67%)	0 (0%)	3 (100%)	0 (0%)	3 (100%)
Sitting to four-point kneeling	3	3 (100%)	0 (0%)	1 (33%)	2 (67%)	1 (33%)	2 (67%)
Sitting without arm support (2)	3	1 (33%)	2 (67%)	2 (67%)	1 (33%)	1 (33%)	2 (67%)
Supported standing (2)	3	3 (100%)	0 (0%)	3 (100%)	0 (0%)	3 (100%)	0 (0%)
Supported standing (3)	-	-	-	-	-	-	-
Pull to stand with support	5	4 (80%)	1 (20%)	1 (20%)	4 (80%)	0 (0%)	5 (100%)
Pulls to stand/stands	1	0 (0%)	1 (100%)	0 (0%)	1 (100%)	0 (0%)	1 (100%)
Supported standing with rotation	4	4 (100%)	0 (0%)	0 (0%)	4 (100%)	0 (0%)	4 (100%)
Cruising without rotation	1	1 (100%)	0 (0%)	0 (0%)	1 (100%)	0 (0%)	1 (100%)
Controlled lowering through standing	4	4 (100%)	0 (0%)	1 (25%)	3 (75%)	1 (25%)	3 (75%)
Cruising with rotation	1	0 (0%)	1 (100%)	0 (0%)	1 (100%)	0 (0%)	1 (100%)
Stands alone	-	-	-	-	-	-	-
Early stepping	-	-	-	-	-	-	-
Standing from modified squat	-	-	-	-	-	-	-
Standing from quadruped position	1	1 (100%)	0 (0%)	1 (100%)	0 (0%)	1 (100%)	0 (0%)
Walks alone	1	1 (100%)	0 (0%)	1 (100%)	0 (0%)	1 (100%)	0 (0%)
Squat	-	-	-	-	-	-	-
38 Movements	**142**	**113 (80%)**	**29 (20%)**	**66 (46%)**	**76 (54%)**	**57 (40%)**	**85 (60%)**

doi:10.6342/NTU202504540

Table 14. Agreement of movement-based assessment results for application version 2.0 to 2.3 using videos with portrait mode and incorrect orientation, cut-off 2025-06-30)

Movements	No. uploaded	Parent vs. PT		AI vs. PT		Parent vs. AI	
		Agreed	Disagreed	Agreed	Disagreed	Agreed	Disagreed
Forearm support (1)	28	10 (36%)	18 (64%)	18 (64%)	0 (36%)	9 (32%)	19 (68%)
Forearm support (2)	8	8 (100%)	0 (0%)	6 (75%)	2 (25%)	6 (75%)	2 (25%)
Extend arm support	14	8 (57%)	6 (43%)	9 (64%)	5 (36%)	5 (36%)	9 (64%)
Reach from forearm support	9	6 (67%)	3 (33%)	4 (44%)	5 (56%)	4 (44%)	5 (56%)
Pivoting	8	6 (75%)	2 (25%)	7 (88%)	1 (13%)	5 (63%)	3 (38%)
Four-point kneeling (1)	13	8 (62%)	5 (38%)	2 (15%)	11 (85%)	2 (15%)	11 (85%)
Reciprocal crawling	10	4 (40%)	6 (60%)	3 (30%)	7 (70%)	4 (40%)	6 (60%)
Four-point kneeling to sitting	12	11 (92%)	1 (8%)	3 (25%)	9 (75%)	3 (25%)	9 (75%)
Reciprocal creeping (1)	6	5 (83%)	1 (17%)	4 (67%)	2 (33%)	4 (67%)	2 (33%)
Reaching from extended arm support	10	9 (90%)	1 (10%)	0 (0%)	10 (100%)	0 (0%)	10 (100%)
Reciprocal creeping (2)	5	4 (80%)	1 (20%)	2 (40%)	3 (60%)	2 (40%)	3 (60%)
Supine lying (3)	20	20 (100%)	0 (0%)	18 (90%)	2 (10%)	18 (90%)	2 (10%)
Supine lying (4)	19	18 (95%)	1 (5%)	13 (68%)	6 (32%)	13 (68%)	6 (32%)
Hands to knees	11	8 (73%)	3 (27%)	8 (73%)	3 (27%)	10 (91%)	1 (9%)
Hands to feet	11	11 (100%)	0 (0%)	11 (100%)	0 (0%)	11 (100%)	0 (0%)
Rolling supine to prone with rotation	12	11 (92%)	1 (8%)	7 (58%)	5 (42%)	8 (67%)	4 (33%)
Sitting with support	30	29 (97%)	1 (3%)	10 (33%)	20 (67%)	11 (37%)	19 (63%)
Sitting with propped arms	22	10 (45%)	12 (55%)	15 (68%)	7 (32%)	9 (41%)	13 (59%)
Pull to sit	9	9 (100%)	0 (0%)	7 (78%)	2 (22%)	7 (78%)	2 (22%)
Sitting with arm support	17	13 (76%)	4 (24%)	10 (59%)	7 (41%)	11 (65%)	6 (35%)
Sitting without arm support (1)	29	22 (76%)	7 (24%)	15 (52%)	14 (48%)	14 (48%)	15 (52%)
Reach with rotation in sitting	15	6 (40%)	9 (60%)	5 (33%)	10 (67%)	3 (20%)	12 (80%)
Sitting to four-point kneeling	13	11 (85%)	2 (15%)	4 (31%)	9 (69%)	4 (31%)	9 (69%)
Sitting without arm support (2)	13	3 (23%)	10 (77%)	6 (46%)	7 (54%)	5 (38%)	8 (62%)
Supported standing (2)	16	16 (100%)	0 (0%)	12 (75%)	4 (25%)	12 (75%)	4 (25%)
Supported standing (3)	8	5 (63%)	3 (38%)	3 (38%)	5 (63%)	3 (38%)	5 (63%)
Pull to stand with support	12	7 (58%)	5 (42%)	4 (33%)	8 (67%)	4 (33%)	8 (67%)
Pulls to stand/stands	16	10 (63%)	6 (38%)	3 (19%)	13 (81%)	3 (19%)	13 (81%)
Supported standing with rotation	12	11 (92%)	1 (8%)	5 (42%)	7 (58%)	5 (42%)	7 (58%)
Cruising without rotation	6	4 (67%)	2 (33%)	2 (33%)	4 (67%)	3 (50%)	3 (50%)
Controlled lowering through standing	7	6 (86%)	1 (14%)	3 (43%)	4 (57%)	3 (43%)	4 (57%)
Cruising with rotation	3	2 (67%)	1 (33%)	1 (33%)	2 (67%)	1 (33%)	2 (67%)
Stands alone	4	4 (100%)	0 (0%)	2 (50%)	2 (50%)	2 (50%)	2 (50%)
Early stepping	3	3 (100%)	0 (0%)	2 (67%)	1 (33%)	2 (67%)	1 (33%)
Standing from modified squat	2	2 (100%)	0 (0%)	0 (0%)	2 (100%)	0 (0%)	2 (100%)
Standing from quadruped position	2	2 (100%)	0 (0%)	2 (100%)	0 (0%)	2 (100%)	0 (0%)
Walks alone	2	2 (100%)	0 (0%)	2 (100%)	0 (0%)	2 (100%)	0 (0%)
Squat	1	1 (100%)	0 (0%)	1 (100%)	0 (0%)	1 (100%)	0 (0%)
38 Movements	**438**	**325 (74%)**	**113 (26%)**	**229 (52%)**	**209 (48%)**	**211 (48%)**	**227 (52%)**

doi:10.6342/NTU202504540

Table 15. Agreement of movement-based assessment results for application version 2.0 to 2.3 using videos with portrait mode or incorrect orientation excluded (Cut-off 2025-06-30)

Movements	No. uploaded	Parent vs. PT		AI vs. PT		Parent vs. AI	
		Agreed	Disagreed	Agreed	Disagreed	Agreed	Disagreed
Forearm support (1)	35	22 (63%)	13 (37%)	25 (71%)	10 (29%)	18 (51%)	17 (49%)
Forearm support (2)	32	25 (78%)	7 (22%)	25 (78%)	7 (22%)	26 (81%)	6 (19%)
Extend arm support	37	18 (49%)	19 (51%)	31 (84%)	6 (16%)	20 (54%)	17 (46%)
Reach from forearm support	20	11 (55%)	9 (45%)	10 (50%)	10 (50%)	9 (45%)	11 (55%)
Pivoting	11	9 (82%)	2 (18%)	10 (91%)	1 (9%)	9 (82%)	2 (18%)
Four-point kneeling (1)	13	10 (77%)	3 (23%)	12 (92%)	1 (8%)	12 (92%)	1 (8%)
Reciprocal crawling	11	5 (45%)	6 (55%)	10 (91%)	1 (9%)	8 (73%)	3 (27%)
Four-point kneeling to sitting	12	11 (92%)	1 (8%)	11 (92%)	1 (8%)	11 (92%)	1 (8%)
Reciprocal creeping (1)	11	9 (82%)	2 (18%)	9 (82%)	2 (18%)	10 (91%)	1 (9%)
Reaching from extended arm support	14	10 (71%)	4 (29%)	4 (29%)	10 (71%)	2 (14%)	12 (86%)
Reciprocal creeping (2)	13	12 (92%)	1 (8%)	11 (85%)	2 (15%)	11 (85%)	2 (15%)
Supine lying (3)	16	12 (75%)	4 (25%)	15 (94%)	1 (6%)	12 (75%)	4 (25%)
Supine lying (4)	22	19 (86%)	3 (14%)	20 (91%)	2 (9%)	19 (86%)	3 (14%)
Hands to knees	16	15 (94%)	1 (6%)	15 (94%)	1 (6%)	15 (94%)	1 (6%)
Hands to feet	22	21 (95%)	1 (5%)	21 (95%)	1 (5%)	21 (95%)	1 (5%)
Rolling supine to prone with rotation	14	14 (100%)	0 (0%)	14 (100%)	0 (0%)	14 (100%)	0 (0%)
Sitting with support	22	22 (100%)	0 (0%)	13 (59%)	9 (41%)	13 (59%)	9 (41%)
Sitting with propped arms	19	13 (68%)	6 (32%)	19 (100%)	0 (0%)	18 (95%)	1 (5%)
Pull to sit	21	20 (95%)	1 (5%)	17 (81%)	4 (19%)	17 (81%)	4 (19%)
Sitting with arm support	21	20 (95%)	1 (5%)	20 (95%)	1 (5%)	19 (90%)	2 (10%)
Sitting without arm support (1)	11	11 (100%)	0 (0%)	10 (91%)	1 (9%)	10 (91%)	1 (9%)
Reach with rotation in sitting	17	11 (65%)	6 (35%)	9 (53%)	8 (47%)	6 (35%)	11 (65%)
Sitting to four-point kneeling	13	10 (77%)	3 (23%)	9 (69%)	4 (31%)	12 (92%)	1 (8%)
Sitting without arm support (2)	10	6 (60%)	4 (40%)	8 (80%)	2 (20%)	8 (80%)	2 (20%)
Supported standing (2)	14	13 (93%)	1 (7%)	12 (86%)	2 (14%)	13 (93%)	1 (7%)
Supported standing (3)	5	5 (100%)	0 (0%)	0 (0%)	5 (100%)	0 (0%)	5 (100%)
Pull to stand with support	13	11 (85%)	2 (15%)	7 (54%)	6 (46%)	7 (54%)	6 (46%)
Pulls to stand/stands	11	10 (91%)	1 (9%)	9 (82%)	2 (18%)	9 (82%)	2 (18%)
Supported standing with rotation	4	2 (50%)	2 (50%)	2 (50%)	2 (50%)	1 (25%)	3 (75%)
Cruising without rotation	6	5 (83%)	1 (17%)	6 (100%)	0 (0%)	6 (100%)	0 (0%)
Controlled lowering through standing	4	3 (75%)	1 (25%)	4 (100%)	0 (0%)	4 (100%)	0 (0%)
Cruising with rotation	1	0 (0%)	1 (100%)	1 (100%)	0 (0%)	0 (0%)	1 (100%)
Stands alone	3	3 (100%)	0 (0%)	2 (67%)	1 (33%)	2 (67%)	1 (33%)
Early stepping	1	1 (100%)	0 (0%)	1 (100%)	0 (0%)	1 (100%)	0 (0%)
Standing from modified squat	1	1 (100%)	0 (0%)	1 (100%)	0 (0%)	1 (100%)	0 (0%)
Standing from quadruped position	1	1 (100%)	0 (0%)	1 (100%)	0 (0%)	1 (100%)	0 (0%)
Walks alone	0	-	-	-	-	-	-
Squat	1	1 (100%)	0 (0%)	1 (100%)	0 (0%)	1 (100%)	0 (0%)
38 Movements	**498**	**392 (79%)**	**106 (21%)**	**395 (79%)**	**103 (21%)**	**366 (73%)**	**132 (27%)**

doi:10.6342/NTU202504540

Table 16. Parental perception for movement-based assessment in preterm and full-term infants using all data with those in portrait mode or incorrect orientation included, (Cut-off 2025-06-30)

Movements	**Preterm (n=543)**			**Full-term (n= 393)**			**Total (n=936)**
	No. uploaded	**Observed**	**Not Observed**	**No. uploaded**	**Observed**	**Not Observed**	
Forearm support (1)	42	23 (55%)	19 (45%)	21	9 (43%)	12 (57%)	63
Forearm support (2)	24	22 (92%)	2 (8%)	16	11 (69%)	5 (31%)	40
Extend arm support	28	14 (50%)	14 (50%)	23	12 (52%)	11 (48%)	51
Reach from forearm support	17	8 (47%)	9 (53%)	12	9 (75%)	3 (25%)	29
Pivoting	8	7 (88%)	1 (13%)	11	8 (73%)	3 (27%)	19
Four-point kneeling (1)	17	12 (71%)	5 (29%)	9	6 (67%)	3 (33%)	26
Reciprocal crawling	12	6 (50%)	6 (50%)	9	3 (33%)	6 (67%)	21
Four-point kneeling to sitting	14	13 (93%)	1 (7%)	10	9 (90%)	1 (10%)	24
Reciprocal creeping (1)	10	9 (90%)	1 (10%)	7	5 (71%)	2 (29%)	17
Reaching from extended arm support	17	14 (82%)	3 (18%)	7	5 (71%)	2 (29%)	24
Reciprocal creeping (2)	9	8 (89%)	1 (11%)	9	8 (89%)	1 (11%)	18
Supine lying (3)	24	20 (83%)	4 (17%)	12	12 (100%)	0 (0%)	36
Supine lying (4)	28	25 (89%)	3 (11%)	13	12 (92%)	1 (8%)	41
Hands to knees	16	14 (88%)	2 (13%)	11	9 (82%)	2 (18%)	27
Hands to feet	19	19 (100%)	0 (0%)	14	13 (93%)	1 (7%)	33
Rolling supine to prone with rotation	15	14 (93%)	1 (7%)	11	11 (100%)	0 (0%)	26
Sitting with support	30	29 (97%)	1 (3%)	22	22 (100%)	0 (0%)	52
Sitting with propped arms	26	13 (50%)	13 (50%)	15	10 (67%)	5 (33%)	41
Pull to sit	16	15 (94%)	1 (6%)	14	14 (100%)	0 (0%)	30
Sitting with arm support	22	19 (86%)	3 (14%)	16	14 (88%)	2 (13%)	38
Sitting without arm support (1)	27	20 (74%)	7 (26%)	13	13 (100%)	0 (0%)	40
Reach with rotation in sitting	15	7 (47%)	8 (53%)	17	10 (59%)	7 (41%)	32
Sitting to four-point kneeling	13	10 (77%)	3 (23%)	13	11 (85%)	2 (15%)	26
Sitting without arm support (2)	12	5 (42%)	7 (58%)	11	4 (36%)	7 (64%)	23
Supported standing (2)	16	15 (94%)	1 (6%)	14	14 (100%)	0 (0%)	30
Supported standing (3)	6	4 (67%)	2 (33%)	7	6 (86%)	1 (14%)	13
Pull to stand with support	16	10 (63%)	6 (38%)	9	8 (89%)	1 (11%)	25
Pulls to stand/stands	12	7 (58%)	5 (42%)	15	13 (87%)	2 (13%)	27
Supported standing with rotation	10	8 (80%)	2 (20%)	6	5 (83%)	1 (17%)	16
Cruising without rotation	6	5 (83%)	1 (17%)	6	4 (67%)	2 (33%)	12
Controlled lowering through standing	5	5 (100%)	0 (0%)	6	4 (67%)	2 (33%)	11
Cruising with rotation	1	1 (100%)	0 (0%)	3	1 (33%)	2 (67%)	4
Stands alone	5	5 (100%)	0 (0%)	2	2 (100%)	0 (0%)	7
Early stepping	1	1 (100%)	0 (0%)	3	3 (100%)	0 (0%)	4
Standing from modified squat	1	1 (100%)	0 (0%)	2	2 (100%)	0 (0%)	3
Standing from quadruped position	1	1 (100%)	0 (0%)	2	2 (100%)	0 (0%)	3
Walks alone	1	1 (100%)	0 (0%)	1	1 (100%)	0 (0%)	2
Squat	1	1 (100%)	0 (0%)	1	1 (100%)	0 (0%)	2
38 Movements	**543**	**411 (76%)**	**132 (24%)**	**393**	**306 (78%)**	**87 (22%)**	**936**

doi:10.6342/NTU202504540

Table 17. Parental perception for movement-based assessment in preterm and full-term infants using all data, with those in portrait mode or incorrect orientation excluded (Cut-off 2025-06-30)

Movements	**Preterm (n=214)**			**Full-term (n= 284)**			**Total (n=498)**
	No. uploaded	**Observed**	**Not Observed**	**No. uploaded**	**Observed**	**Not Observed**	
Forearm support (1)	19	14 (74%)	5 (26%)	16	8 (50%)	8 (50%)	35
Forearm support (2)	19	17 (89%)	2 (11%)	13	8 (62%)	5 (38%)	32
Extend arm support	17	7 (41%)	10 (59%)	20	11 (55%)	9 (45%)	37
Reach from forearm support	8	2 (25%)	6 (75%)	12	9 (75%)	3 (25%)	20
Pivoting	1	1 (100%)	0 (0%)	10	8 (80%)	2 (20%)	11
Four-point kneeling (1)	4	4 (100%)	0 (0%)	9	6 (67%)	3 (33%)	13
Reciprocal crawling	3	2 (67%)	1 (33%)	8	3 (38%)	5 (63%)	11
Four-point kneeling to sitting	4	4 (100%)	0 (0%)	8	7 (88%)	1 (13%)	12
Reciprocal creeping (1)	5	5 (100%)	0 (0%)	6	4 (67%)	2 (33%)	11
Reaching from extended arm support	9	7 (78%)	2 (22%)	5	3 (60%)	2 (40%)	14
Reciprocal creeping (2)	5	5 (100%)	0 (0%)	8	7 (88%)	1 (13%)	13
Supine lying (3)	7	3 (43%)	4 (57%)	9	9 (100%)	0 (0%)	16
Supine lying (4)	13	11 (85%)	2 (15%)	9	8 (89%)	1 (11%)	22
Hands to knees	9	8 (89%)	1 (11%)	7	7 (100%)	0 (0%)	16
Hands to feet	9	9 (100%)	0 (0%)	13	12 (92%)	1 (8%)	22
Rolling supine to prone with rotation	6	6 (100%)	0 (0%)	8	8 (100%)	0 (0%)	14
Sitting with support	11	11 (100%)	0 (0%)	11	11 (100%)	0 (0%)	22
Sitting with propped arms	7	4 (57%)	3 (43%)	12	9 (75%)	3 (25%)	19
Pull to sit	10	9 (90%)	1 (10%)	11	11 (100%)	0 (0%)	21
Sitting with arm support	9	9 (100%)	0 (0%)	12	11 (92%)	1 (8%)	21
Sitting without arm support (1)	2	2 (100%)	0 (0%)	9	9 (100%)	0 (0%)	11
Reach with rotation in sitting	4	3 (75%)	1 (25%)	13	8 (62%)	5 (38%)	17
Sitting to four-point kneeling	2	1 (50%)	1 (50%)	11	9 (82%)	2 (18%)	13
Sitting without arm support (2)	4	3 (75%)	1 (25%)	6	3 (50%)	3 (50%)	10
Supported standing (2)	5	4 (80%)	1 (20%)	9	9 (100%)	0 (0%)	14
Supported standing (3)	2	2 (100%)	0 (0%)	3	3 (100%)	0 (0%)	5
Pull to stand with support	7	6 (86%)	1 (14%)	6	5 (83%)	1 (17%)	13
Pulls to stand/stands	5	4 (80%)	1 (20%)	6	6 (100%)	0 (0%)	11
Supported standing with rotation	1	0 (0%)	1 (100%)	3	2 (67%)	1 (33%)	4
Cruising without rotation	3	3 (100%)	0 (0%)	3	2 (67%)	1 (33%)	6
Controlled lowering through standing	2	2 (100%)	0 (0%)	2	1 (50%)	1 (50%)	4
Cruising with rotation	0	-	-	1	0 (0%)	1 (100%)	1
Stands alone	2	2 (100%)	0 (0%)	1	1 (100%)	0 (0%)	3
Early stepping	0	-	-	1	1 (100%)	0 (0%)	1
Standing from modified squat	0	-	-	1	1 (100%)	0 (0%)	1
Standing from quadruped position	0	-	-	1	1 (100%)	0 (0%)	1
Walks alone	0	-	-	0	–	–	0
Squat	0	-	-	1	1 (100%)	0 (0%)	1
38 Movements	**214**	**170 (79%)**	**44 (21%)**	**284**	**222 (78%)**	**62 (22%)**	**498**

doi:10.6342/NTU202504540

Table 18. Parental perception of movement-based assessment in preterm and full-term infants using data from Google form (Time duration: 2025-03-09 to 2025-06-30)

Movements	Preterm (n=129)			Full-term (n= 70)			Total (n=202)
	No. uploaded	Observed	Not Observed	No. uploaded	Observed	Not Observed	
Forearm support (1)	6	5 (83%)	1 (17%)	0	0 (0%)	0 (0%)	6
Forearm support (2)	4	4 (100%)	0 (0%)	2	2 (100%)	0 (0%)	6
Extend arm support	2	2 (100%)	0 (0%)	5	4 (80%)	1 (20%)	7
Reach from forearm support	3	3 (100%)	0 (0%)	2	2 (100%)	0 (0%)	5
Pivoting	8	8 (100%)	0 (0%)	3	3 (100%)	0 (0%)	11
Four-point kneeling (1)	2	2 (100%)	0 (0%)	1	1 (100%)	0 (0%)	3
Reciprocal crawling	1	1 (100%)	0 (0%)	2	2 (100%)	0 (0%)	3
Four-point kneeling to sitting	2	2 (100%)	0 (0%)	1	1 (100%)	0 (0%)	3
Reciprocal creeping (1)	2	2 (100%)	0 (0%)	2	2 (100%)	0 (0%)	4
Reaching from extended arm support	4	3 (75%)	1 (25%)	2	2 (100%)	0 (0%)	5
Reciprocal creeping (2)	2	2 (100%)	0 (0%)	0	0 (0%)	0 (0%)	2
Supine lying (3)	3	3 (100%)	0 (0%)	0	0 (0%)	0 (0%)	3
Supine lying (4)	0	0 (0%)	0 (0%)	0	0 (0%)	0 (0%)	0
Hands to knees	3	2 (67%)	1 (33%)	2	2 (100%)	0 (0%)	5
Hands to feet	7	6 (86%)	1 (14%)	7	6 (86%)	1 (14%)	14
Rolling supine to prone with rotation	15	15 (100%)	0 (0%)	6	6 (100%)	0 (0%)	21
Sitting with support	3	3 (100%)	0 (0%)	0	0 (0%)	0 (0%)	3
Sitting with propped arms	2	2 (100%)	0 (0%)	2	2 (100%)	0 (0%)	4
Pull to sit	2	2 (100%)	0 (0%)	2	2 (100%)	0 (0%)	4
Sitting with arm support	3	3 (100%)	0 (0%)	4	4 (100%)	0 (0%)	6
Sitting without arm support (1)	8	7 (87%)	1 (13%)	1	1 (80%)	0 (0%)	12
Reach with rotation in sitting	5	4 (80%)	1 (20%)	2	2 (100%)	0 (0%)	8
Sitting to four-point kneeling	2	2 (100%)	0 (0%)	1	1 (100%)	0 (0%)	3
Sitting without arm support (2)	7	7 (100%)	0 (0%)	3	3 (100%)	0 (0%)	10
Supported standing (2)	14	14 (100%)	0 (0%)	12	12 (100%)	0 (0%)	26
Supported standing (3)	2	2 (100%)	0 (0%)	1	1 (100%)	0 (0%)	3
Pull to stand with support	2	2 (100%)	0 (0%)	2	1 (50%)	1 (50%)	3
Pulls to stand/stands	2	2 (100%)	0 (0%)	1	1 (100%)	0 (0%)	3
Supported standing with rotation	2	2 (100%)	0 (0%)	1	1 (100%)	0 (0%)	3
Cruising without rotation	4	4 (100%)	0 (0%)	2	2 (100%)	0 (0%)	6
Controlled lowering through standing	2	2 (100%)	0 (0%)	0	0 (0%)	0 (0%)	3
Cruising with rotation	2	2 (100%)	0 (0%)	1	1 (100%)	0 (0%)	3
Stands alone	2	1 (50%)	1 (50%)	1	1 (100%)	0 (0%)	3
Early stepping	3	2 (67%)	1 (33%)	0	0 (0%)	0 (0%)	3
Standing from modified squat	0	0 (0%)	0 (0%)	0	0 (0%)	0 (0%)	0
Standing from quadruped position	0	0 (0%)	0 (0%)	0	0 (0%)	0 (0%)	0
Walks alone	0	0 (0%)	0 (0%)	0	0 (0%)	0 (0%)	0
Squat	0	0 (0%)	0 (0%)	0	0 (0%)	0 (0%)	0
38 Movements	**131**	**123 (94%)**	**8 (6%)**	**71**	**68 (96%)**	**3 (4%)**	**202**

doi:10.6342/NTU202504540

Table 19. Agreements between the parents', AI, and physiotherapists' results for age-based assessment using completed data from the application, with those in portrait and incorrect orientation included (Cut-off 2025-06-30)

Age (months)	No. of uploaded trials	Parent vs. PT		AI vs. PT		Parent vs. AI	
		All Movements Observed	Partial Movements Observed	All Movements Observed	Partial Movements Observed	All Movements Observed	Partial Movements Observed
3	7	7 (100%)	0 (0%)	5 (71%)	2 (29%)	3 (43%)	4 (57%)
4	13	12 (92%)	1 (8%)	10 (77%)	3 (23%)	9 (69%)	4 (31%)
5	15	14 (93%)	1 (7%)	12 (80%)	3 (20%)	11 (73%)	4 (27%)
6	10	9 (90%)	1 (10%)	8 (80%)	2 (20%)	8 (80%)	2 (20%)
7	11	10 (91%)	1 (9%)	8 (73%)	3 (27%)	6 (55%)	5 (45%)
8	8	8 (100%)	0 (0%)	5 (63%)	3 (38%)	4 (50%)	4 (50%)
9	6	6 (100%)	0 (0%)	6 (100%)	0 (0%)	4 (67%)	2 (33%)
10	5	5 (100%)	0 (0%)	4 (80%)	1 (20%)	3 (60%)	2 (40%)
11	3	3 (100%)	0 (0%)	2 (67%)	1 (33%)	1 (33%)	2 (67%)
12	1	1 (100%)	0 (0%)	1 (100%)	0 (0%)	1 (100%)	0 (0%)
13-18	1	1 (100%)	0 (0%)	1 (100%)	0 (0%)	1 (100%)	0 (0%)
3-18 months	**80**	**76 (95%)**	**4 (5%)**	**62 (78%)**	**18 (23%)**	**51 (64%)**	**29 (36%)**

doi:10.6342/NTU202504540

Table 20. Agreements between the parents', AI, and physiotherapists' results for age-based assessment using completed data from the application, with those in portrait and incorrect orientation excluded (Cut-off 2025-06-30)

Age-Group	No. uploaded	Parent vs. PT		AI vs. PT		Parent vs. AI	
		All Movements Observed	Partial Movements Observed	All Movements Observed	Partial Movements Observed	All Movements Observed	Partial Movements Observed
3	3	3 (100%)	0 (0%)	3 (100%)	0 (0%)	3 (100%)	0 (0%)
4	11	11 (100%)	0 (0%)	10 (91%)	1 (9%)	7 (64%)	4 (36%)
5	11	11 (100%)	0 (0%)	10 (91%)	1 (9%)	9 (82%)	2 (18%)
6	9	8 (89%)	1 (11%)	7 (78%)	2 (22%)	8 (89%)	1 (11%)
7	10	9 (90%)	1 (10%)	7 (70%)	3 (30%)	6 (60%)	4 (40%)
8	6	6 (100%)	0 (0%)	3 (50%)	3 (50%)	4 (67%)	2 (33%)
9	4	4 (100%)	0 (0%)	4 (100%)	0 (0%)	3 (75%)	1 (25%)
10	3	3 (100%)	0 (0%)	3 (100%)	0 (0%)	2 (67%)	1 (33%)
11	2	2 (100%)	0 (0%)	1 (50%)	1 (50%)	1 (50%)	1 (50%)
12	1	1 (100%)	0 (0%)	1 (100%)	0 (0%)	1 (100%)	0 (0%)
13 - 18	1	1 (100%)	0 (0%)	1 (100%)	0 (0%)	1 (100%)	0 (0%)
3-18 months	**61**	**59 (97%)**	**2 (3%)**	**50 (82%)**	**11 (18%)**	**45 (74%)**	**16 (26%)**

doi:10.6342/NTU202504540

Table 21. Agreements between the parents', AI, and physiotherapists' results for age-based assessment in preterm and full-term infants using completed data from the application, with those in portrait and incorrect orientation included (Cut-off 2025-06-30)

Age-Group	Preterm			Full-term			Total Number
	Number	All Movements Observed	Partial Movements Observed	Number	All Movements Observed	Partial Movements Observed	
3	4	4 (100%)	0 (0%)	3	3 (100%)	0 (0%)	7
4	9	8 (89%)	1 (11%)	4	4 (100%)	0 (0%)	13
5	7	6 (86%)	1 (14%)	7	7 (100%)	0 (0%)	15
6	4	4 (100%)	0 (0%)	6	5 (83%)	1 (17%)	10
7	5	5 (100%)	0 (0%)	7	6 (86%)	1 (14%)	11
8	2	2 (100%)	0 (0%)	5	5 (100%)	0 (0%)	8
9	2	2 (100%)	0 (0%)	4	4 (100%)	0 (0%)	6
10	3	3 (100%)	0 (0%)	4	4 (100%)	0 (0%)	5
11	4	4 (100%)	0 (0%)	0	0 (0%)	0 (0%)	3
12	1	1 (100%)	0 (0%)	0	0 (0%)	0 (0%)	1
13 - 18	0	0 (0%)	0 (0%)	1	1 (100%)	0 (0%)	1
3-18 months	**41**	**39 (95%)**	**2 (5%)**	**40**	**38 (95%)**	**2 (5%)**	**81**

doi:10.6342/NTU202504540

Table 22. Agreements between the parents', AI, and physiotherapists' results for age-based assessment in preterm and full-term infants using completed data from the application, with those in portrait and incorrect orientation excluded (Cut-off 2025-06-30)

Age-Group	**Preterm**			**Full-term**			**Total Number**
	Number	**All Movements Observed**	**Partial Movements Observed**	**Number**	**All Movements Observed**	**Partial Movements Observed**	
3	3	3 (100%)	0 (0%)	0	0 (0%)	0 (0%)	3
4	7	7 (100%)	0 (0%)	4	4 (100%)	0 (0%)	11
5	5	4 (80%)	1 (20%)	6	6 (100%)	0 (0%)	11
6	3	3 (100%)	0 (0%)	6	5 (83%)	1 (17%)	9
7	3	3 (100%)	0 (0%)	7	6 (86%)	1 (14%)	10
8	0	0 (0%)	0 (0%)	6	6 (100%)	0 (0%)	6
9	0	0 (0%)	0 (0%)	4	4 (100%)	0 (0%)	4
10	1	1 (100%)	0 (0%)	2	2 (100%)	0 (0%)	3
11	2	2 (100%)	0 (0%)	0	0 (0%)	0 (0%)	2
12	1	1 (100%)	0 (0%)	0	0 (0%)	0 (0%)	1
13 - 18	0	0 (0%)	0 (0%)	1	1 (100%)	0 (0%)	1
3-18 months	**25**	**24 (96%)**	**1 (4%)**	**36**	**34 (94%)**	**2 (6%)**	**61**

doi:10.6342/NTU202504540

Table 23. Agreements between the parents, AI, and physiotherapists' results for age-based assessment in preterm and full-term infants using data from Google form (Time duration: 2025-03-09 to 2025-06-30)

Age (Month)	Preterm (n=33)			Full-term (n=17)			Total (n=50)		
	No. uploaded	All Movements Observed	Partial Movements Observed	No. uploaded	All Movements Observed	Partial Movements Observed	No. uploaded	All Movements Observed	Partial Movements Observed
3	2	2 (100%)	0 (0%)	-	-	-	2	2 (100%)	0 (0%)
4	2	2 (100%)	0 (0%)	-	-	-	2	2 (100%)	0 (0%)
5	3	2 (67%)	1 (33%)	2	1 (50%)	1 (50%)	5	3 (60%)	2 (40%)
6	3	3 (100%)	0 (0%)	4	3 (75%)	1 (25%)	7	6 (75%)	1 (25%)
7	3	3 (100%)	0 (0%)	2	2 (100%)	0 (0%)	5	5 (100%)	0 (0%)
8	9	7 (78%)	2 (12%)	3	2 (67%)	1 (33%)	12	9 (75%)	3 (25%)
9	1	1 (100%)	0 (0%)	2	2 (100%)	0 (0%)	3	3 (100%)	0 (0%)
10	1	1 (100%)	0 (0%)	2	2 (100%)	0 (0%)	3	3 (100%)	0 (0%)
11	4	3 (75%)	1 (25%)	1	1 (100%)	0 (0%)	5	4 (80%)	1 (20%)
12	2	2 (100%)	0 (0%)	-	-	-	2	2 (100%)	0 (0%)
13	1	0 (0%)	1 (100%)	1	1 (100%)	0 (0%)	2	1 (50%)	1 (50%)
14	1	0 (0%)	1 (100%)	-	-	-	1	0 (0%)	1 (100%)
15	1	0 (0%)	1 (100%)	-	-	-	1	0 (0%)	1 (100%)
16	-	-	-	-	-	-	-	-	-
17		-	-	-	-	-	-	-	-
18		-	-	-	-	-	-	-	-
3-18 months	**33**	**26 (79%)**	**7 (21%)**	**17**	**14 (82%)**	**3 (18%)**	**50**	**40 (80%)**	**10 (20%)**

doi:10.6342/NTU202504540

Table 24. Validity of the parental perception with PT results (Cut-off 2025-06-30)

Age in Months	No. of Infants	Movement Set[a]	PT Results[b]		Agreement	SEN	SPE	PPV	NPV
			Not within the expected range	Within the expected range					
3	2	Not within the expected range	1	0	1.00	1.00	1.00	1.00	1.00
		Within the expected range	0	1					
4	2	Not within the expected range	0	0	1.00	-	1.00	-	1.00
		Within the expected range	0	2					
5	5	Not within the expected range	1	0	1.00	1.00	1.00	1.00	1.00
		Within the expected range	0	4					
6	9	Not within the expected range	1	0	0.88	0.50	1.00	1.00	0.88
		Within the expected range	1	7					
7	8	Not within the expected range	2	0	1.00	1.00	1.00	1.00	1.00
		Within the expected range	0	6					
8	12	Not within the expected range	4	1	1.00	1.00	0.88	0.80	1.00
		Within the expected range	0	7					
9	5	Not within the expected range	0	0	0.80	-	1.00	-	0.80
		Within the expected range	0	5					
10	4	Not within the expected range	0	0	1.00	-	1.00	-	1.00
		Within the expected range	0	4					
11	5	Not within the expected range	0	1	0.83	-	0.83	-	1.00
		Within the expected range	0	5					
12	2	Not within the expected range	0	0	1.00	-	1.00	-	1.00
		Within the expected range	0	2					
13 – 18	6	Not within the expected range	3	0	1.00	1.00	1.00	1.00	1.00
		Within the expected range	0	3					
TOTAL	**61**	Not within the expected range	**12**	**2**	**0.93**	**0.86**	**0.96**	**0.86**	**0.96**
		Within the expected range	**2**	**45**					

SEN = Sensitivity, SPE = Specificity, PPV = Positive predictive value, NPV = Negative predictive value
Within the expected range was defined as the infant showing presence of all movements appropriate for the assessment age (all movements observed).
Not within the expected range was defined as the absence of one or more movements within the normal age movement criteria (partial movements observed).

doi:10.6342/NTU202504540

Table 25. Validation of AI results with the onsite AIMS assessment (Cut-off 2025-06-30)

Subject	Age (Months)	Group	Parent-AI Result	AIMS Score	AIMS Percentile	AIMS Result	Interpretation
118-1	4	Preterm	Normal	17	50-75%	Normal	True Negative
118-2	4	Preterm	Normal	16	50-75%	Normal	True Negative
159	5	Preterm	Normal	20	25-50%	Normal	True Negative
148	7	Preterm	Normal	28	10-25%	Normal	True Negative
124	12	Preterm	Normal	52	25-50%	Normal	True Negative
124	16	Preterm	Normal	58	> 90%	Normal	True Negative
125	8	Preterm	Delay	39	75-90%	Normal	False Positive
117	6	Full-term	Normal	20	25-50%	Normal	True Negative
117	7	Full-term	Normal	30	25-50%	Normal	True Negative
120	8	Full-term	Normal	46	> 90%	Normal	True Negative
124	14	Preterm	Delay	52	> 5%	Delay	True Positive
131	3	Full-term	Delay	4	> 5%	Delay	True Positive
126	6	Full-term	Delay	22	10-25%	Normal	False Positive
128	13	Full-term	Delay	54	25-50%	Normal	False Positive

doi:10.6342/NTU202504540

Table 26. Concurrent validity of AI results with the onsite AIMS assessment (Cut-off 2025-06-30)

Age in Months	No. of Infants	Movement Set[a]	PT Results[b] Not within the expected range	PT Results[b] Within the expected range	Agreement	SEN	SPE	PPV	NPV
3	1	Not within the expected range	1	0	1.00	1.00	-	1.00	-
		Within the expected range	0	0					
4	2	Not within the expected range	0	0	1.00	-	1.00	-	1.00
		Within the expected range	0	2					
5	1	Not within the expected range	0	0	1.00	-	1.00	-	1.00
		Within the expected range	0	1					
6	2	Not within the expected range	0	1	0.50	-	0.50	0.00	1.00
		Within the expected range	0	1					
7	2	Not within the expected range	0	0	1.00	-	1.00	-	1.00
		Within the expected range	0	2					
8	2	Not within the expected range	0	1	0.50	-	0.50	0.00	1.00
		Within the expected range	0	1					
12	1	Not within the expected range	0	0	1.00	-	1.00	-	0.80
		Within the expected range	0	1					
13	1	Not within the expected range	0	1	1.00	-	1.00	-	1.00
		Within the expected range	0	0					
14	1	Not within the expected range	1	0	1.00	1.00	-	1.00	-
		Within the expected range	0	0					
16	1	Not within the expected range	0	0	1.00	-	1.00	-	1.00
		Within the expected range	0	1					
TOTAL	**14**	Not within the expected range	**2**	**3**	**0.79**	**1.00**	**0.75**	**0.40**	**1.00**
		Within the expected range	**0**	**9**					

SEN = Sensitivity, SPE = Specificity, PPV = Positive predictive value, NPV = Negative predictive value
Within the expected range was defined as the infant showing presence of all movements appropriate for the assessment age (all movements observed).
Not within the expected range was defined as the absence of one or more movements within the normal age movement criteria (partial movements observed).

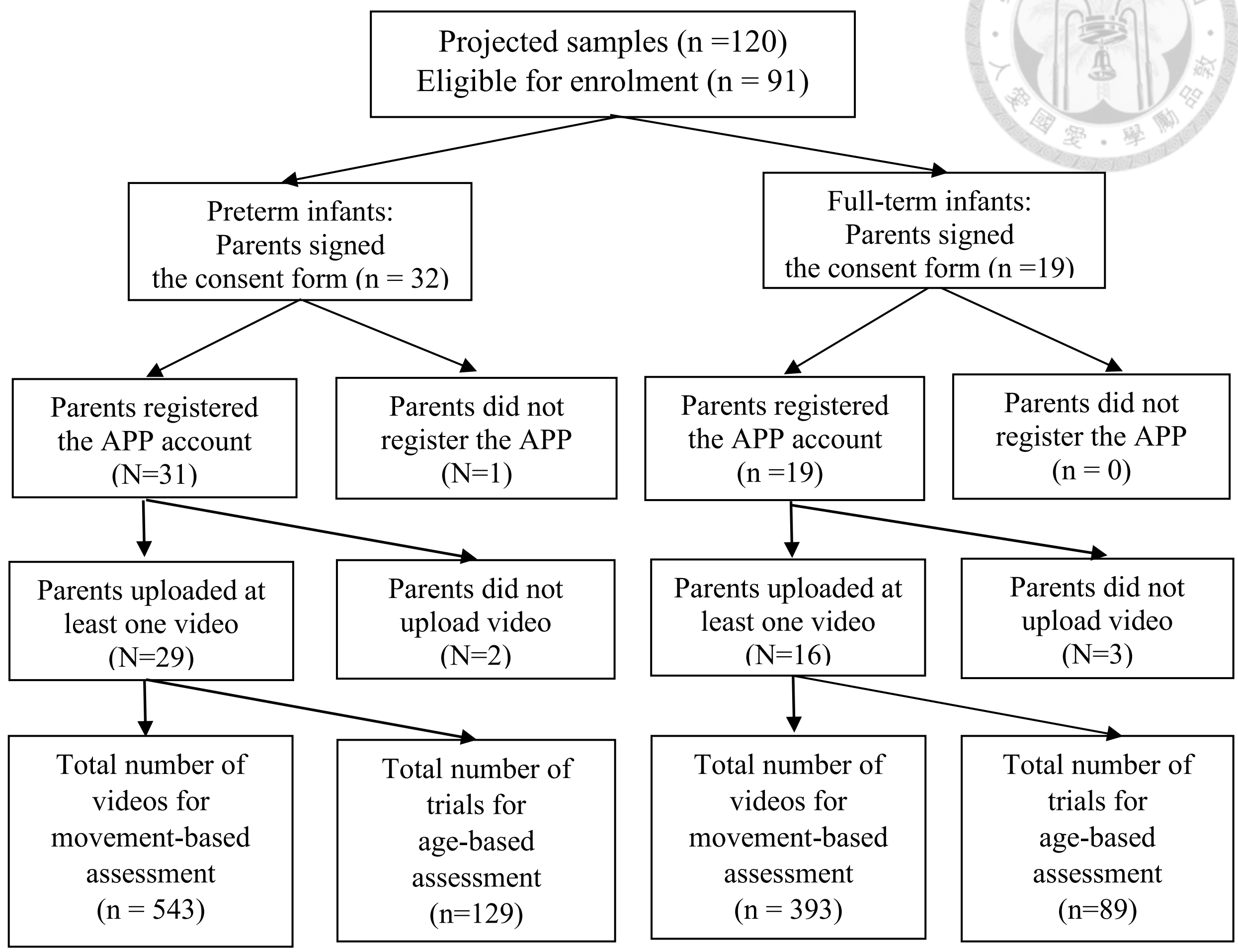
doi:10.6342/NTU202504540

Table 27. Preliminary study of parental perception assessment compared with AIMS assessment conducted by physiotherapists

Age	Number	Agreement (%)	Disagreement (%)	Item of disagreement	Reason of disagreement
4	5	3 (60%)	2 (40%)	- Supported standing (2) (n=2)	- Parent difficult to do this item alone while recording - Parenting culture (don't put in standing in early months of age)
6	3	3 (100%)	0 (0%)	-	-
8	3	3 (100%)	0 (0%)	-	-
10	3	1 (33%)	2 (67%)	- Sitting without arm support (2) (n=1) - Supported standing with rotation (n=2)	- Parent misperception about the leg's movement - Parent didn't notice the rotation of the trunk
12	5	5 (100%)	0 (0%)	-	-
13	1	1 (100%)	0 (0%)	-	-
14	2	2 (100%)	0 (0%)	-	-
TOTAL	**22**	**18 (82%)**	**4 (18%)**		

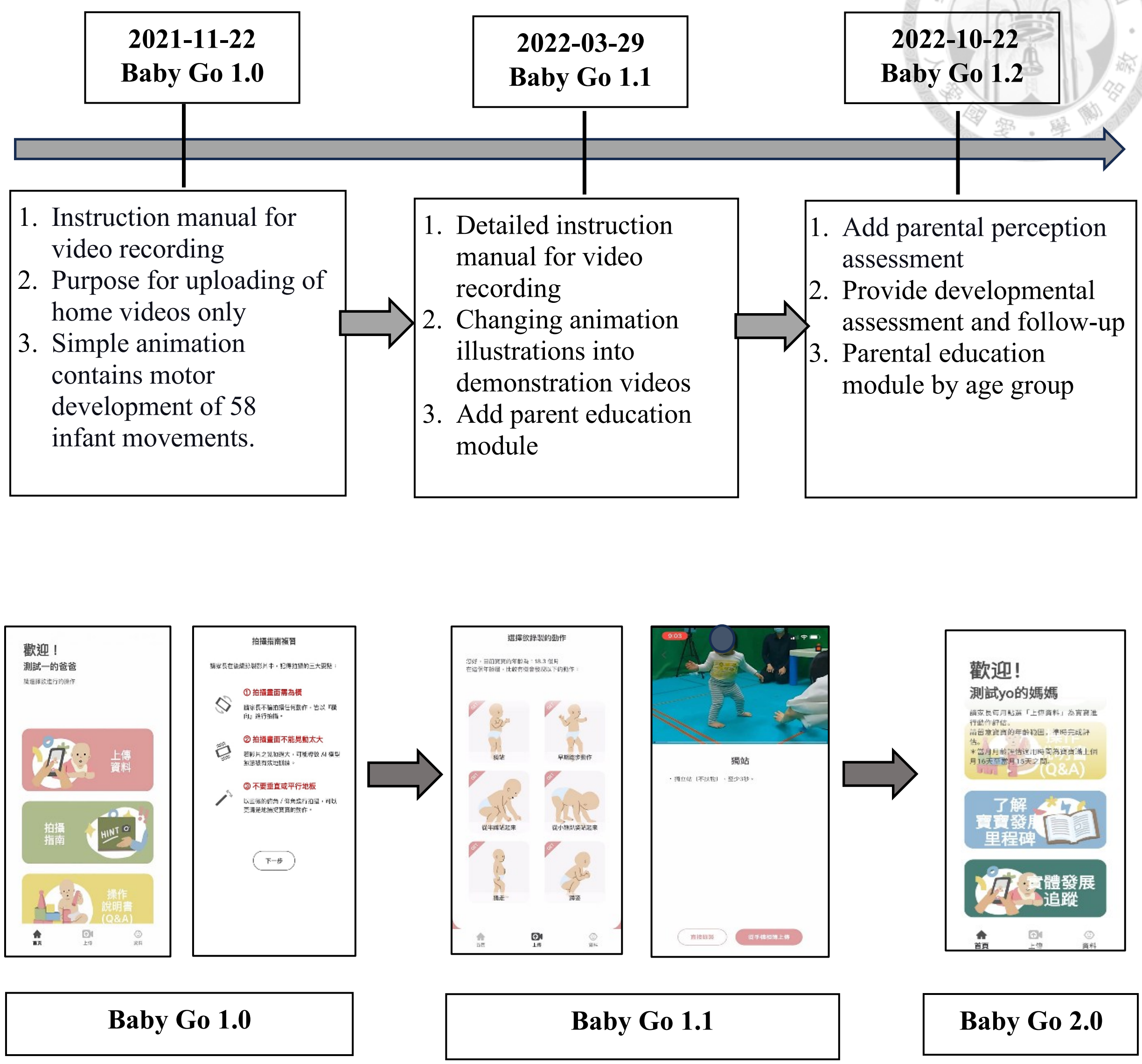
doi:10.6342/NTU202504540

Figure 1. **Data collection flow chart** (Cut-off on 2025-06-30)



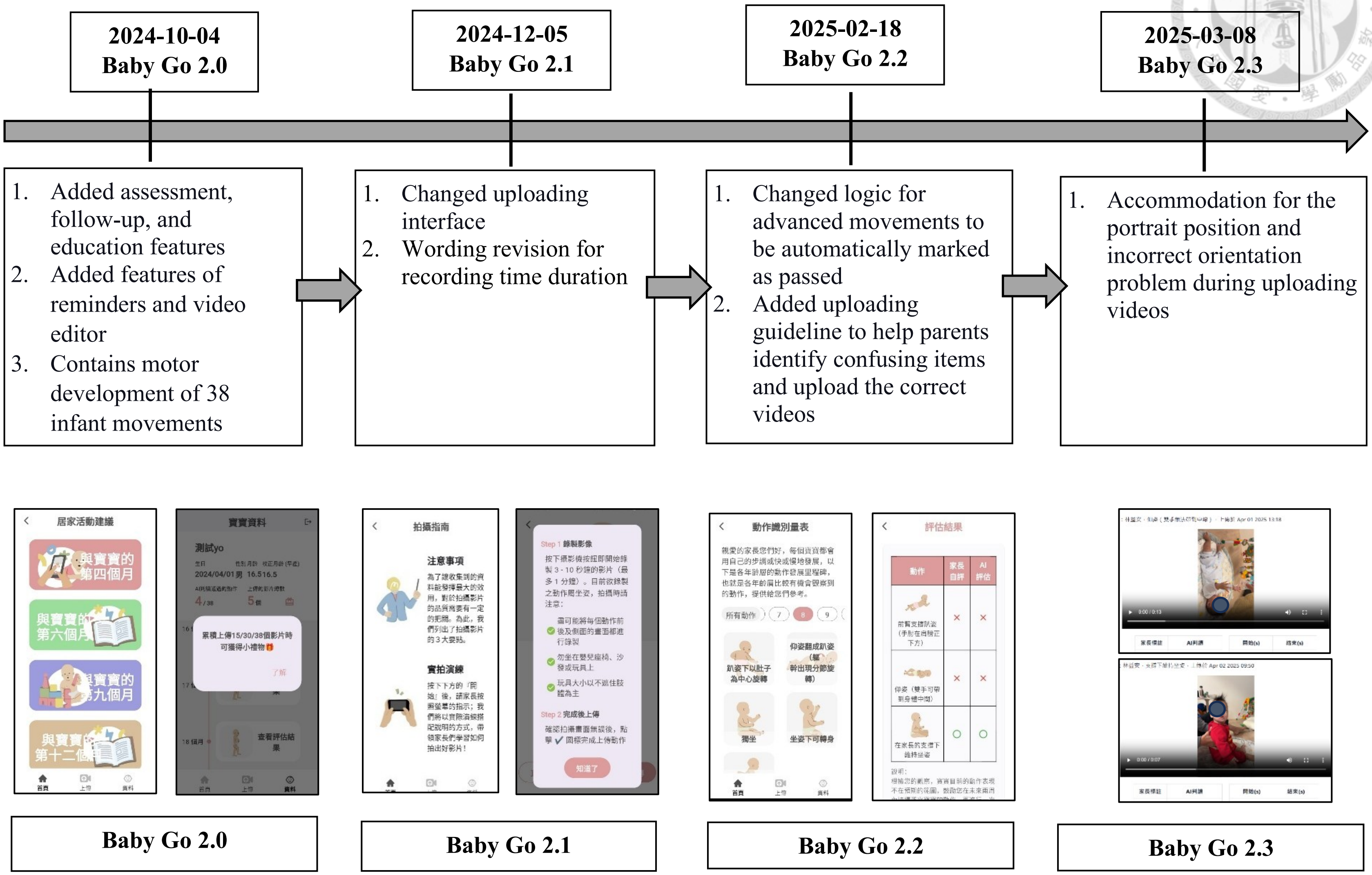
doi:10.6342/NTU202504540

Figure 2. Timeline for development and revision of the "*Baby Go*" app from version 1.0 to 1.2 in the previous study (2021-11-22 – 2022-10-22)



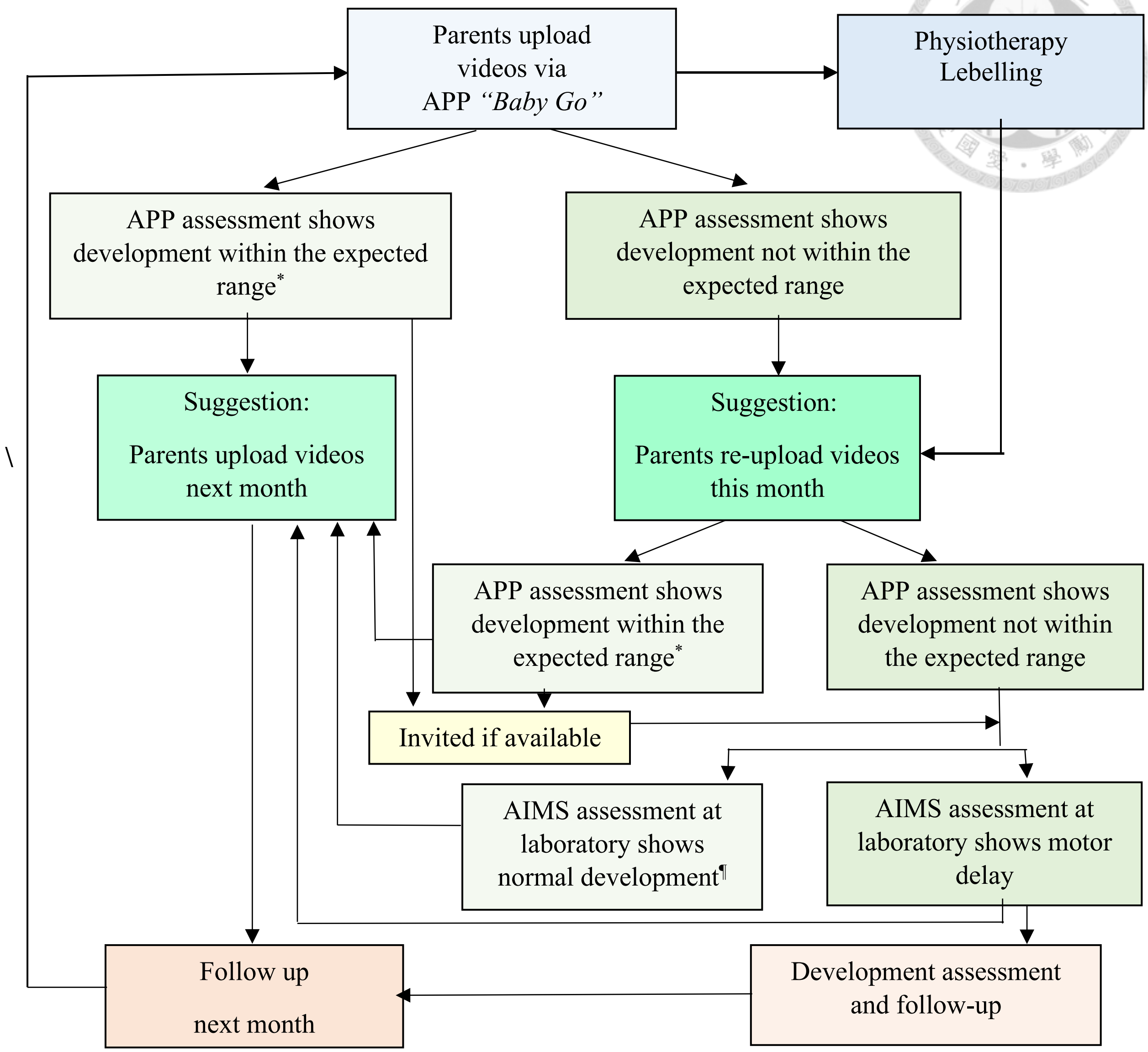
doi:10.6342/NTU202504540

Figure 3. Timeline for the development and revision of the "*Baby Go*" app from version 2.0 to 2.3 in this study (2024-10-04 to 2025-03-08)



doi:10.6342/NTU202504540

Figure 4. **Flow chart for APP developmental assessment and follow-up**



AIMS = Alberta Infant Motor Scale.
* "Development within the expected range" in the APP is defined as presence of all items at assessment age.
¶ "Normal development" in the AIMS assessment is defined as 10^{th}-90^{th} % at assessment age.

doi:10.6342/NTU202504540

Appendices

Appendix 1. Movements in the parental perception assessment

Prone Movements (n=11)	Supine Movements (n=5)	Sitting Movements (n=8)	Standing Movements (n=14)
Prone lying (1)	Supine lying (1)	Sitting with support	Supported standing (1)
Prone lying (2)	Supine lying (2)	Sitting with propped arms	Supported standing (2)
Prone prop	Supine lying (3)	Pull to sit	Supported standing (3)
Forearm support (1)	Supine lying (4)	Unsustainable sitting	Pull to stand with support
Prone mobility	Hands to knees	Sitting with arm support	Pulls to stand/stands
Forearm support (2)	Active extension	Unsustainable sitting without arm support	Supported standing with rotation
Extend arm support	Hands to feet	Weight shift in unsustained sitting	Cruising without rotation
Rolling to supine without rotation	Rolling supine to prone without rotation	Sitting without arm support (1)	Half-kneeling
Swimming	Rolling supine to prone with rotation	Reach with rotation in sitting	Controlled lowering through standing
Reach from forearm support		Sitting to prone	Cruising with rotation
Pivoting		Sitting to four-point kneeling	Stands alone
Rolling prone to supine with rotation		Sitting without arm support (2)	Early stepping
Four-point kneeling (1)			Standing from modified squat
Propped side lying			Standing from quadruped position
Reciprocal crawling			Walks alone
Four-point kneeling to sitting			Squat
Reciprocal creeping (1)			
Reaching from extended arm support			
Four-point kneeling (2)			
Modified four-point kneeling			
Reciprocal creeping (2)			

doi:10.6342/NTU202504540

Appendix 2. Illustration of the distribution of 38 movements

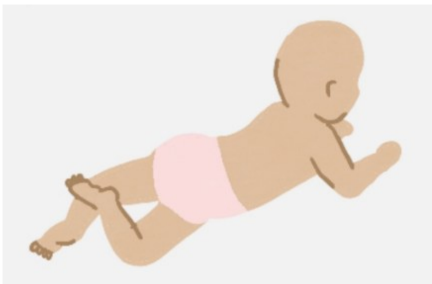
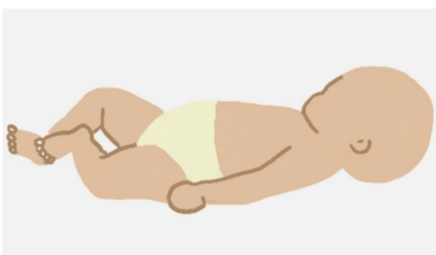
Age (Mon)	**Prone (N=11)**	**Supine (N=5)**	**Sitting (N=8)**	**Standing (N=14)**
3	Forearm support (1)	Supine lying (3)	Sitting with support	
4	Forearm support (1)	Supine lying (4)	Sitting with support	
5	Forearm support (2)	Hands to knees	Sitting with propped arms Pull to sit	Supported standing (2)
6	Extended arm support	Hands to feet	Sitting with arm support	Supported standing (2)
7	Reaching from arm support	Hands to feet	Sitting with arm support	Supported standing (2)
8	Pivoting	Rolling to prone with rotation	Sitting without arm support (1) Reach with rotation in sitting	Supported standing (2)
9	Four points kneeling (1)	Rolling to prone with rotation	Sitting to four points kneeling	Supported standing (3) Pull to stand with support
10	Reciprocal crawling, Four points kneeling to sitting		Sitting without arm support (2)	Pull to stand/stance Supported standing with rotation
11	Reciprocal creeping (1) Reaching from extended arm support		Sitting without arm support (2)	Cruising without rotation
12	Reciprocal creeping (2)		Sitting without arm support (2)	Cruising without rotation Controlled lowering through standing
13	Reciprocal creeping (2)			Cruising with rotation Stands alone
14				Stands alone Early stepping
15				Standing from quadruped position Standing from modified squat Walks alone
16-18				Walks alone Squatting

doi:10.6342/NTU202504540

Appendix 3. Parental Perception Assessment Checklist

三 個 月

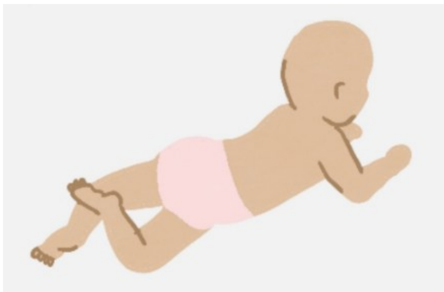
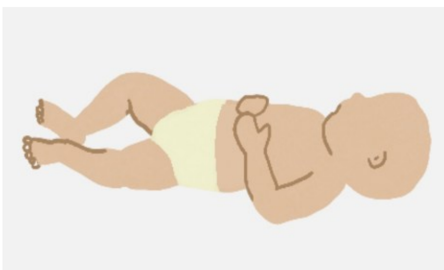
ID:　　　　　　姓名：

	**前臂支撐趴姿（手肘在肩膀正下方）** 手肘撐著地板趴著，胸部離開地面。 頭部可以抬高到與地面夾角超過 45 度， 並維持在身體中間。 從側面看，寶寶的手肘位在肩膀的正下方或前方。 **家 長 分 數 ： □ 有出現 □ 沒出現** **P T 分 數 ： □ 有出現 □ 沒出現**
	**仰姿（雙手無法帶到身體中間）** 仰躺時，頭可以維持在身體中線（身體的中間）， 雙手在身體兩側揮動，雙腳踢動。 **家 長 分 數 ： □ 有出現 □ 沒出現** **P T 分 數 ： □ 有出現 □ 沒出現**
	**在家長的支撐下維持坐姿** 家長雙手支撐在寶寶腋下時，寶寶可以維持坐姿， 並讓頭與身體呈現一直線。 **家 長 分 數 ： □ 有出現 □ 沒出現** **P T 分 數 ： □ 有出現 □ 沒出現**

doi:10.6342/NTU202504540

四 個 月

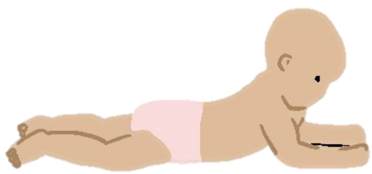
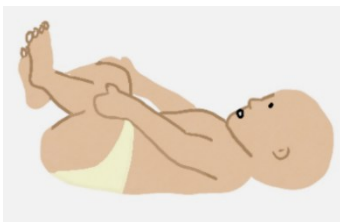
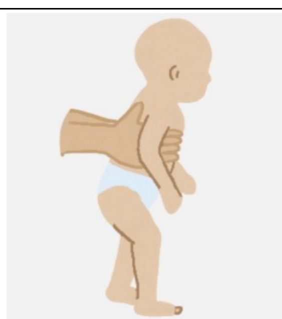
ID: 姓名：

	**前臂支撐趴姿（手肘在肩膀正下方）** 手肘撐著地板趴著，胸部離開地面。 頭部可以抬高到與地面夾角超過 45 度， 並維持在身體中間。 從側面看，寶寶的手肘位在肩膀的正下方或前方。 **家長分數：☐ 有出現 ☐ 沒出現** **PT 分數：☐ 有出現 ☐ 沒出現**
	**仰姿（雙手可帶到身體中間）** 仰躺時，頭可以維持在身體中線（身體的中間）， 雙手在胸前玩玩具，雙腳踢動。 **家長分數：☐ 有出現 ☐ 沒出現** **PT 分數：☐ 有出現 ☐ 沒出現**
	**在家長的支撐下維持坐姿** 家長雙手支撐在寶寶腋下時，寶寶可以維持坐姿， 並讓頭與身體呈現一直線。 **家長分數：☐ 有出現 ☐ 沒出現** **PT 分數：☐ 有出現 ☐ 沒出現**

doi:10.6342/NTU202504540

五 個 月

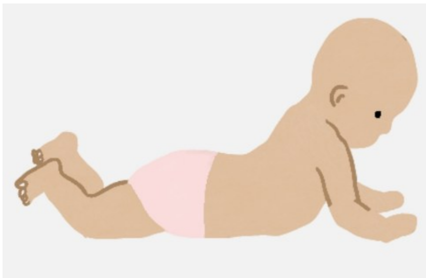
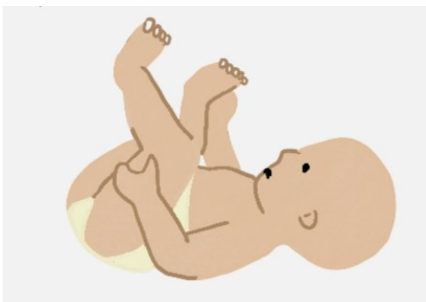
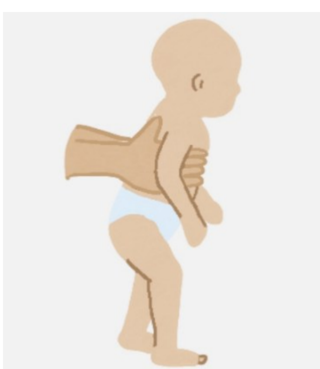
ID:　　　　　姓名：

	**前臂支撐趴姿（手肘在肩膀前）** 手肘撐地趴著，胸部離地。 頭部可以抬高到與地面夾角 90 度，會出現縮下巴動作。 從側面看，寶寶的手肘位在肩膀的前方。 **家長分數：☐ 有出現 ☐ 沒出現** **PT 分數：☐ 有出現 ☐ 沒出現**
	**手摸膝蓋** 仰躺時，膝蓋彎起，手碰觸到膝蓋。 **家長分數：☐ 有出現 ☐ 沒出現** **PT 分數：☐ 有出現 ☐ 沒出現**
	**身體前傾用手幫忙支撐維持坐姿** 家長將寶寶擺放在坐姿，寶寶身體會前傾， 頭部維持在中線，雙手向前撐地， 自己維持坐姿至少 3 秒。 **家長分數：☐ 有出現 ☐ 沒出現** **PT 分數：☐ 有出現 ☐ 沒出現**
	**仰姿拉成坐姿** 寶寶在仰躺姿勢，家長以聲音或目光吸引寶寶的注意 ，並握住寶寶的手腕拉成坐姿。 過程中寶寶的頭部及身體需維持一直線。 **家長分數：☐ 有出現 ☐ 沒出現** **PT 分數：☐ 有出現 ☐ 沒出現**
	**支撐下站立（可控制頭部）** 在家長雙手扶住寶寶兒腋下站立時， 寶寶可以將頭抬起與身體平行。 雙腳微微承重，下肢可能出現伸直、踩踏、或踢腳動作。 從側面觀察，屁股的位置會落在肩膀後方。 **家長分數：☐ 有出現 ☐ 沒出現** **PT 分數：☐ 有出現 ☐ 沒出現**

doi:10.6342/NTU202504540

六 個 月

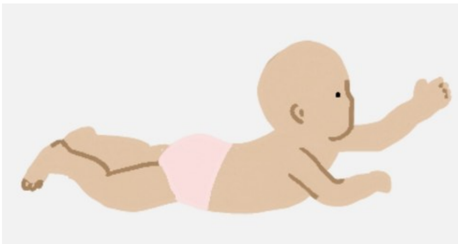
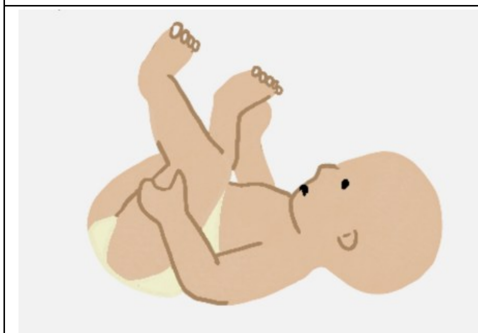
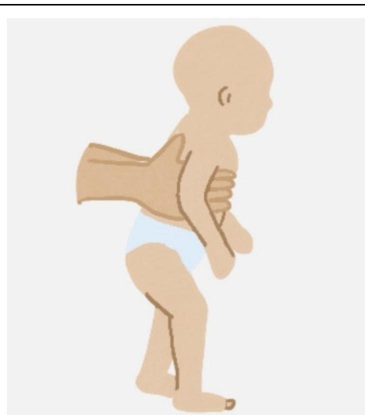
ID:　　　　　姓名：

	**手掌支撐的趴** 手掌撐地趴著，胸部離地。 **家長分數：□有出現□沒出現** **PT 分數：□有出現□沒出現**
	**手摸腳腳** 仰躺時，將腳舉起到靠近肚子或胸口的地方。手可以觸碰腳或是抓腳玩。 **家長分數：□有出現□沒出現** **PT 分數：□有出現□沒出現**
	**用手支撐維持坐姿** 家長將寶寶擺放在坐姿，寶寶可以雙手伸直撐地坐著。身體挺直，頭部可以自由地左右看。無法變換姿勢或移動重心。 **家長分數：□有出現□沒出現** **PT 分數：□有出現□沒出現**
	**支撐下站立（可控制頭部）** 在家長雙手扶住寶寶兒腋下站立時， 寶寶可以將頭抬起與身體平行。 雙腳微微承重，下肢可能出現伸直、踩踏、或踢腳動作。 從側面觀察，屁股的位置會落在肩膀後方。 **家長分數：□有出現□沒出現** **PT 分數：□有出現□沒出現**

doi:10.6342/NTU202504540

七 個 月

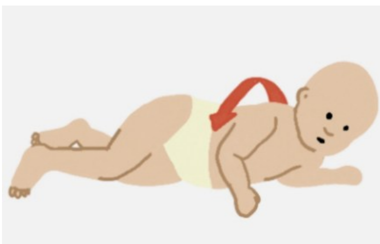
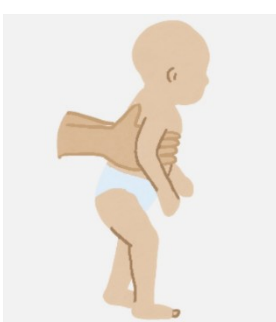
ID:　　　　　姓名：

	**在手肘撐地時伸手取物** 手肘撐地趴著，伸一隻手向前向上(至少到眼睛高度) 觸碰或拿取玩具。 **家長分數：□ 有出現 □ 沒出現** **PT 分數：□ 有出現 □ 沒出現**
	**手摸腳腳** 仰躺時，將腳舉起到靠近肚子或胸口的地方。 手可以觸碰腳或是抓腳玩。 **家長分數：□ 有出現 □ 沒出現** **PT 分數：□ 有出現 □ 沒出現**
	**用手支撐維持坐姿** 家長將寶寶擺放在坐姿，寶寶可以雙手伸直撐地坐著。 身體挺直，頭部可以自由地左右看。無法變換姿勢或移動重心。 **家長分數：□ 有出現 □ 沒出現** **PT 分數：□ 有出現 □ 沒出現**
	**支撐下站立（可控制頭部）** 在家長雙手扶住寶寶兒腋下站立時， 寶寶可以將頭抬起與身體平行。 雙腳微微承重，下肢可能出現伸直、踩踏、或踢腳動作。 從側面觀察，屁股的位置會落在肩膀後方。 **家長分數：□ 有出現 □ 沒出現** **PT 分數：□ 有出現 □ 沒出現**

doi:10.6342/NTU202504540

八 個 月

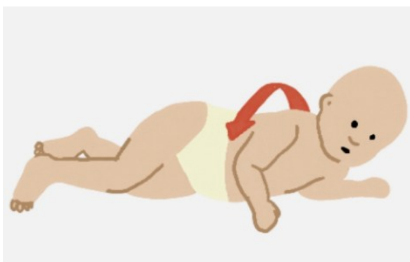
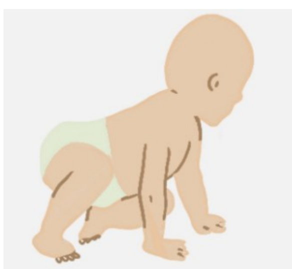
ID:　　　　　姓名：

	**趴姿下以肚子為中心旋轉** 趴姿下，以肚子為中心旋轉。 寶寶會用手撐地向左或向右移動。 **家 長 分 數 ： ☐ 有出現 ☐ 沒出現** **P T　分 數 ： ☐ 有出現 ☐ 沒出現**
	**仰姿翻成趴姿（軀幹出現分節旋轉）** 從仰躺翻身變成趴姿，可能由頭、肩膀或是腳先啟動。身體呈現分節式翻轉。 **家 長 分 數 ： ☐ 有出現 ☐ 沒出現** **P T　分 數 ： ☐ 有出現 ☐ 沒出現**
	**獨坐** 寶寶可以獨立坐著玩一段時間。 手可以離開身體玩玩具，家長不需要擔心坐不穩而倒下。 **家 長 分 數 ： ☐ 有出現 ☐ 沒出現** **P T　分 數 ： ☐ 有出現 ☐ 沒出現**
	**坐姿下可轉身** 寶寶獨坐時，可以向後轉身去去摸家長從寶寶側後方給的玩具，不會因為轉身而失去平衡。 **家 長 分 數 ： ☐ 有出現 ☐ 沒出現** **P T　分 數 ： ☐ 有出現 ☐ 沒出現**
	**支撐下站立（可控制頭部）** 在家長雙手扶住寶寶兒腋下站立時， 寶寶可以將頭抬起與身體平行。 雙腳微微承重，下肢可能出現伸直、踩踏、或踢腳動作。 從側面觀察，屁股的位置會落在肩膀後方。 **家 長 分 數 ： ☐ 有出現 ☐ 沒出現** **P T　分 數 ： ☐ 有出現 ☐ 沒出現**

doi:10.6342/NTU202504540

九 個 月

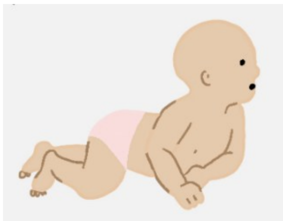
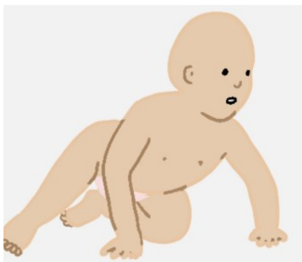
ID: 姓名：

	**四點跪姿（背部不平）** 維持雙手手掌、兩側膝蓋著地的四點跪姿。 肚子離開地面但是下垂、背部向地面凹陷。 **家 長 分 數 ： ☐ 有出現 ☐ 沒出現** **P T 分 數 ： ☐ 有出現 ☐ 沒出現**
	**仰姿翻成趴姿（軀幹出現分節旋轉）** 從仰躺翻身變成趴姿，可能由頭、肩膀或是腳先啟動。身體呈現分節式翻轉。 **家 長 分 數 ： ☐ 有出現 ☐ 沒出現** **P T 分 數 ： ☐ 有出現 ☐ 沒出現**
	**坐姿變四點跪姿** 變換動作的過程中，運用兩腳的力量將屁股抬起，由坐姿變成四點跪姿。 **家 長 分 數 ： ☐ 有出現 ☐ 沒出現** **P T 分 數 ： ☐ 有出現 ☐ 沒出現**
	**支撐下站立（身體成一直線）** 在家長雙手扶住寶寶腋下站立時， 寶寶可以將頭抬起與身體平行。 雙腳出現起立蹲下、單腳站、雙腳伸直等動作。 從側面看，屁股的位置與肩膀成一直線。 **家 長 分 數 ： ☐ 有出現 ☐ 沒出現** **P T 分 數 ： ☐ 有出現 ☐ 沒出現**
	**雙手扶物從坐到站** 自己拉著家具的邊緣從坐到站。 家長可將玩具放置在小椅子或桌子上引導孩童做出動作。 **家 長 分 數 ： ☐ 有出現 ☐ 沒出現** **P T 分 數 ： ☐ 有出現 ☐ 沒出現**

doi:10.6342/NTU202504540

十 個 月

ID: 姓名：

	**肚子貼地爬行** 肚子貼在地板上，以匍匐的方式向前或向後爬行。 **家 長 分 數 ：☐ 有出現 ☐ 沒出現** **P T 分 數 ：☐ 有出現 ☐ 沒出現**
	**四點跪姿變成坐姿** 從雙手手掌、兩側膝蓋著地的四點跪姿，變換姿勢成坐姿。 **家 長 分 數 ：☐ 有出現 ☐ 沒出現** **P T 分 數 ：☐ 有出現 ☐ 沒出現**
	**獨坐（可變換不同坐姿）** 可以自由變換不同的坐姿，雙腳可以擺成多樣化的姿勢。可以輕鬆的重心轉移、變換姿勢。 **家 長 分 數 ：☐ 有出現 ☐ 沒出現** **P T 分 數 ：☐ 有出現 ☐ 沒出現**
	**雙手扶物從坐到站，可轉移重心** 自己拉著家具的邊緣從坐到站。 扶站時雙腳可以踩得很穩，且可以靈活轉換重心 家長可以將玩具放置在小椅子或桌子上來引導孩童做出動作。 **家 長 分 數 ：☐ 有出現 ☐ 沒出現** **P T 分 數 ：☐ 有出現 ☐ 沒出現**
	**雙手扶物站立，可轉身** 扶著家具站立時，可以轉身放開一手去拿身體側後方的玩具。 **家 長 分 數 ：☐ 有出現 ☐ 沒出現** **P T 分 數 ：☐ 有出現 ☐ 沒出現**

doi:10.6342/NTU202504540

十一 個 月

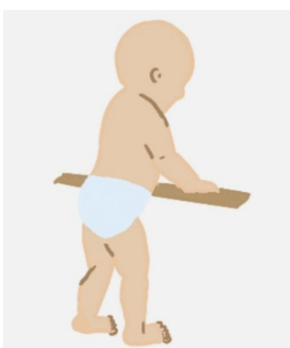
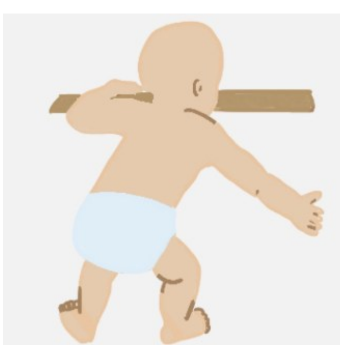
ID:　　　　　姓名：

	**肚子離地爬行（背部不平）** 肚子離地往前爬。 肚子下垂、背部向地面凹陷。 **家長分數：□ 有出現 □ 沒出現** **PT 分數：□ 有出現 □ 沒出現**
	**四點跪姿下伸手取物** 在雙手手掌、兩側膝蓋著地的四點跪姿下。 伸一隻手向前向上(至少到眼睛高度)觸碰或拿取玩具。 **家長分數：□ 有出現 □ 沒出現** **PT 分數：□ 有出現 □ 沒出現**
	**獨坐（可變換不同坐姿）** 可以自由變換不同的坐姿，雙腳可以擺成多樣化的姿勢。 可以輕鬆的重心轉移、變換姿勢。 **家長分數：□ 有出現 □ 沒出現** **PT 分數：□ 有出現 □ 沒出現**
	**扶物側走，無法轉身** 扶著家具側走，身體跟桌面保持平行， 沒有出現轉身的動作。 **家長分數：□ 有出現 □ 沒出現** **PT 分數：□ 有出現 □ 沒出現**

doi:10.6342/NTU202504540

十二 個 月

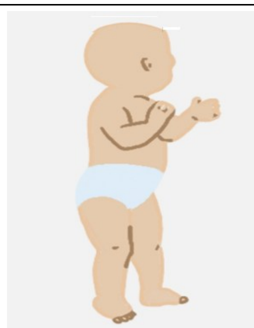
ID:　　　　姓名：

	**肚子離地爬行（背部平直）** 肚子離地往前爬。 背部平直、肚子沒有下垂。 **家長分數：□有出現□沒出現** **PT 分數：□有出現□沒出現**
	**獨坐（可變換不同坐姿）** 可以自由變換不同的坐姿，雙腳可以擺成多樣化的姿勢。 可以輕鬆的重心轉移、變換姿勢。 **家長分數：□有出現□沒出現** **PT 分數：□有出現□沒出現**
	**扶物側走，無法轉身** 扶著家具側走，身體跟桌面保持平行， 沒有出現轉身的動作。 **家長分數：□有出現□沒出現** **PT 分數：□有出現□沒出現**
	**手扶物蹲下（慢慢將重心放低）** 扶物站立時，一手扶物、另一手往地面摸/拿東西。 雙腳慢慢蹲下將重心放低，過程中不會跌坐在地上。 **家長分數：□有出現□沒出現** **PT 分數：□有出現□沒出現**

doi:10.6342/NTU202504540

十三 個 月

ID:　　　　　　姓名：

	**肚子離地爬行（背部平直）** 肚子離地往前爬。 背部平直、肚子沒有下垂。 **家 長 分 數 ：☐ 有出現 ☐ 沒出現** **P T 分 數 ：☐ 有出現 ☐ 沒出現**
	**扶物側走，可轉身** 扶物側走時伴隨轉身。 **家 長 分 數 ：☐ 有出現 ☐ 沒出現** **P T 分 數 ：☐ 有出現 ☐ 沒出現**
	**獨站** 放手站，至少 3 秒。 **家 長 分 數 ：☐ 有出現 ☐ 沒出現** **P T 分 數 ：☐ 有出現 ☐ 沒出現**

doi:10.6342/NTU202504540

十四 個 月

ID:　　　　　　姓名：

	**獨站** 放手站，至少 3 秒。 **家長分數：□ 有出現 □ 沒出現** **PT 分數：□ 有出現 □ 沒出現**
	**早期踏步動作** 放手連續走至少 5 步。 因為寶寶剛學會走路，此時腳步快速、步伐短，雙手會抬高到胸口高度幫助自己維持平衡。 家長需在周圍維護安全避免跌倒。 **家長分數：□ 有出現 □ 沒出現** **PT 分數：□ 有出現 □ 沒出現**

doi:10.6342/NTU202504540