www.nature.com/pr

CLINICAL RESEARCH ARTICLE

Caregiver-reported newborn term and preterm motor abilities: psychometrics of the PediaTrac^{TM} Motor domain

Renee Lajiness-O'Neill[1,2]✉, Trivellore Raghunathan[3], Patricia Berglund[3], Alissa Huth-Bocks[4], H. Gerry Taylor[5], Angela D. Staples[1], Judith Brooks[6], Angela Lukomski[7], Jennifer C. Gidley Larson[2], Seth Warschausky[2] and PediaTrac Project Consortium*

BACKGROUND: Approximately 5–10% of children exhibit developmental deviations in motor skills or other domains; however, physicians detect less than one-third of these abnormalities. Systematic tracking and early identification of motor deviations are fundamental for timely intervention.
METHODS: Term and preterm neonates were prospectively assessed at the newborn (NB) period in a study of the psychometric properties of the Motor (MOT) domain of PediaTrac^{TM} v3.0, a novel caregiver-based development tracking instrument. Item response theory graded response modeling was used to model item parameters and estimate theta, an index of the latent trait, motor ability. Exploratory factor analysis (EFA) was conducted to examine the dimensionality and factor structure.
RESULTS: In a cohort of 571 caregiver/infant dyads (331 term, 240 preterm), NB MOT domain reliability was high (rho = 0.94). Item discrimination and item difficulty of each of the 15 items could be reliably modeled across the range of motor ability. EFA confirmed that the items constituted a single dimension with second-order factors, accounting for 43.20% of variance.
CONCLUSIONS: The latent trait, motor ability, could be reliably estimated at the NB period.

Pediatric Research (2023) 93:1736–1744; https://doi.org/10.1038/s41390-022-02312-4

IMPACT:

- The caregiver-reported Motor domain of PediaTrac provides a reliable estimate of the latent trait of motor ability during the newborn period.
- This is the first known caregiver-reported instrument that can assess motor ability in the newborn period with high reliability in term and preterm infants.
- Item response theory methods were employed that will allow for future characterization of developmental subgroups and motor trajectories.
- The PediaTrac Motor domain can support early identification of at-risk infants.
- Including caregivers in digital reporting and child-centered monitoring of motor functioning may improve access to care.

INTRODUCTION

During early development, motor, socioemotional, communication, and cognitive abilities develop concurrently and mutually influence each other.[1] Adverse early biological events such as prematurity, genetic or neurodevelopmental disorders can have negative and cascading effects on development.[2–5] Given how rapidly development progresses in infancy and the interdependency of these early skills, it is crucial to identify developmental deviations at the time of onset, before they negatively affect subsequent stages of development. The sooner a developmental deviation is detected, the earlier intervention can occur.[6] Surveillance, risk detection, and early identification are vital steps toward ensuring that children receive timely services.[7]

Systematic and prospective developmental tracking

Systematically tracking developmental domains beyond anthropometric growth parameters may be a practical method for monitoring such risk.[8] Consistent with this hypothesis, the Centers for Disease Control and Prevention's (CDC) "Learn the Signs. Act Early." program recently convened an expert working group to modify its developmental surveillance checklists to identify evidence-based milestones for inclusion in CDC checklists.[9] However, checklists are not scales per se and cannot provide a measurement of developmental trajectories. Tracking development in a manner similar to tracking physical growth parameters may provide a more refined method for early detection of risk. For example, concurrent tracking of trajectories of motor and

[1]Psychology, Eastern Michigan University, Ypsilanti, MI, USA. [2]Physical Medicine and Rehabilitation, University of Michigan, Ann Arbor, MI, USA. [3]Institute for Social Research and the School of Public Health, University of Michigan, Ann Arbor, MI, USA. [4]Pediatrics, Rainbow Babies & Children's Hospital, Case Western Reserve University, Cleveland, OH, USA. [5]Abigail Wexner Research Institute at Nationwide Children's Hospital, and Pediatrics, The Ohio State University, Columbus, OH, USA. [6]Dietetics and Human Nutrition, Eastern Michigan University, Ypsilanti, MI, USA. [7]Nursing, Eastern Michigan University, Ypsilanti, MI, USA. *A list of authors and their affiliations appears at the end of the paper.
✉email: rlajines@emich.edu

Received: 15 November 2021 Revised: 18 August 2022 Accepted: 7 September 2022
Published online: 30 September 2022

communication milestones in infancy may identify children at risk for autism spectrum disorder.[10]

Although approximately 5–10% of children exhibit developmental disorders that manifest in deficits in skills and/or behavior, it is estimated that primary care physicians detect less than one-third of these delays.[11–13] In part, this may be due to the lack of a systematic method to collect and synthesize information about development from caregivers. Infants born preterm (<37 weeks gestation) are at a particularly high risk for developmental delays.[14] Direct assessment methods such as term-age magnetic resonance imaging, the Prechtl Qualitative Assessment of General Movements, and the Hammersmith Infant Neurological Examination are important tools for detecting risk for CP and other neurodevelopmental conditions in infants who had neonatal complications such as preterm birth.[15,16] However, these sensitive instruments are less likely to be employed with infants who did not have significant neonatal complications. Additional direct assessment methods of motor and other developmental domains are utilized, but these methods are resource intensive, expensive, and typically reserved for infants at known risk. Both the Bayley Scales of Infant and Toddler Development and Mullen Scales of Early Learning have gross and fine motor domains, and these measures have strong psychometric properties.[17,18] However, these tools require extensive expertise, and many geographic regions do not have access to qualified professionals to conduct such assessments. In the United States, approximately 16 million children live in areas where there is no or limited access to a local pediatrician.[17,19] Relying exclusively on direct and time-intensive assessment methods may contribute to disparities in access to care in less-resourced regions.

Comprehensive caregiver report measures

As an alternative to direct measurement methods, caregiver report methods are not resource-prohibitive.[20,21] A very limited number of comprehensive self-administered caregiver report measures exist for infants such as the Ages & Stages Questionnaires, Third Edition (ASQ-3) system.[22] In a recent review of the psychometric quality of instruments that allow for a multidimensional assessment of child development from 0 to 12 years, only seven instruments were identified.[23] Five of those instruments required direct assessment by an examiner and only two could be completed independently as caregiver report, the Vineland Adaptive Behavior Scales (VABS) and the ASQ.[24] The VABS focuses on adaptive behavior and does not allow for serial assessment or a method to track trajectories. Of the instruments investigated, the ASQ-3 was suggested as the most appropriate to screen child development in epidemiological studies or large-scale evaluation in health services. However, there is limited information reported about the internal structure or external validity of the ASQ-3. The ASQ uses pre-established cutoffs to identify risk, rather than examining deviations from developmental trajectories per se. The DAYC-2, which assesses both fine and gross motor abilities in children aged 0–5 years within its Physical Development Domain, can be based on caregiver report but only when administered by a trained expert via caregiver interview, and it is typically used as a direct assessment measure by qualified clinicians.[25] None of the industry standard caregiver reports or for-profit diagnostic management system efforts, such as the Child Health and Development Interactive System (CHADIS) or Patient-Reported Outcomes Measurement Information System (PROMIS), offers a method to systematically track trajectories of motor development, and neither CHADIS nor PROMIS assess motor development across infancy.[26,27]

Development of a caregiver-reported measure using item response theory

To continuously track motor trajectories, a test construction method such as item response theory (IRT) must be employed. IRT is a measurement framework that employs latent trait methods to model psychometric functioning at both the item and test level.[28,29] IRT uses mathematical models that explain the relationship between latent traits (attributes) and their observed outcomes. IRT models the likelihood of a given response to an item as a probabilistic function of the individual's score on a latent trait of interest, referred to as theta.[30] In this investigation, the latent trait is motor functioning. Reliability is reframed as measurement precision, and is represented in IRT models using the concept of "information."[31] Theta values are computed across ability, for both items and the full test. IRT offers benefits over classical test theory including sample-invariant parameter estimates (i.e., assuming no differential item functioning across populations) to metrics of reliability at both the item and test level.[28,29,32]

The initial step in the application of IRT is to estimate the parameters that describe the relationship between the item and the individual. In IRT, it is necessary to assess the fit of the model to the data. IRT items and ability parameters are said to be invariant, as ability estimates obtained from different sets of items will be the same, and item parameter estimates (explained further below) obtained in different groups of examinees will be the same. IRT is an item-oriented rather than a test-oriented test construction method. As such, it lends itself to an individualized medicine approach in assessment and subsequent care.

This psychometric study examines the measurement precision and validation of the newborn (NB) Motor (MOT) domain of PediaTrac^{TM}, a web-based survey tool involving prospective data collection throughout infancy across multiple domains of development via caregiver report. It was hypothesized that (1) parameter estimates of the items would be reliably modeled, (2) the latent trait, motor ability, would be reliably estimated by theta, with estimates of the reliability of MOT in the acceptable to good range (>0.70), and (3) exploratory factor analysis (EFA) would provide support for both a single primary motor dimension and possible secondary factors indicating distinct types of motor skills.

METHODS

Participants

This investigation is part of a larger prospective, longitudinal investigation, the PediaTrac study, of a sample of 571 caregivers of infants (48% female) who were born either at term ($n = 331$; 49% female) or preterm ($n = 240$; 46% female).[21] The sample was recruited from three sites that included academic medical centers and a local community center: 100 from Site #1, 239 from Site #2, and 232 from Site #3. Term infants had a gestational age (GA) of ≥37 weeks at birth and minimum birth weight of 2500 g, with no history of prenatal or intrapartum complications, neonatal abstinence syndrome, neurological injury/disease, or known genetic disorder. Preterm infants had a GA of <37 weeks. Birth weight was allowed to vary, but exclusions from the preterm group included neonatal abstinence syndrome, neurological injury or disease unrelated to preterm birth, and Down syndrome. For multiple births, randomization was used to enroll infants. Caregivers were a minimum of 18 years old and had access to a personal device such as a smartphone, tablet, or computer. Ninety-eight percent of the respondents were biological mothers. English-language competence was required for participation. All American Psychological Association ethical guidelines were followed and a Reliant Institutional Review Board approval was obtained. See Table 1 for participant characteristics and descriptive summaries for the pooled sample.

Study procedures

Women were recruited in their last trimester of pregnancy, after their infants' birth in the hospital, or at their first NB visit, with consent obtained after birth. Primary caregivers of term infants completed PediaTrac soon after birth, whereas caregivers of preterm infants completed it when their infants reached a postmenstrual age of 39 weeks. All subsequent data collection timepoints were based on corrected age for preterm infants.

Table 1. Pooled demographics and characteristics of infants and caregivers.

Infant sex	
Male	298 (52.2%)
Female	273 (47.8%)
Term status	
Full-term	331 (58.0%)
Preterm	240 (42.0%)
Gestational age at birth (weeks)	
Full-term	39.0 ± 1.15
Preterm	33.0 ± 2.96
Self-identified infant race[a]	
Black or African American	192 (34.1%)
Multiracial	58 (10.3%)
Other	5 (0.89%)
White	272 (48.3%)
Ethnicity	
Hispanic/Latino	36 (6.39%)
Maternal age at enrollment (years)	30.1 ± 6.04
Household ADI	5.32 ± 3.33
Household income	
Below poverty	169 (32.8%)
Below median	70 (13.6%)
At/above median	122 (23.7%)
At/above twice median	88 (17.1%)
Above $150,000	66 (12.8%)
Maternal education	
Some/completed high school	133 (23.3%)
Some college/trade, technical, or vocational training	160 (28.0%)
College graduate	136 (23.8%)
Some/completed postgraduate or professional degree	142 (24.9%)
Caregiver marital status[b]	
Married	305 (53.5%)
Not married	265 (46.5%)

Gestational age at birth, maternal age, and household ADI are presented as mean ± standard deviation. Income was categorized relative to the US Department of Health and Human Services Poverty Guidelines (2019) and median household income for (state blinded for review). It should be noted that the difference in median household income in 2019 for (other state blinded for review), though smaller, was similar enough to (blinded) that the categorization would be the same whether blinded state or blinded state was used. This categorization is based on the number of people in the home as well as household income. The total number of cases differs, for example, from the total number of participants because of missing income information either from declining to state $n = 46$ or no response $n = 10$.

ADI Area Deprivation Index, using state deciles.

[a]Race categories White, Black or African American, Multiracial, and Other are Non-Hispanic. Participants self-identified as Multiracial or Other. Race was unknown for 8 infants.

[b]A total of 98% of caregivers were biological mothers, and marital status was missing for 1 caregiver.

Study measure and variables

PediaTrac v3.0 is a web-based survey comprising between 511 and 558 unique items covering the age range from birth to 18 months.[21] Caregivers complete subsets of the PediaTrac survey ranging from ~220 to 340 items, depending on the time period of the assessment. PediaTrac queries multiple developmental domains (Feeding/Eating/Elimination, Sleep, Motor, Social/Communication/Cognition, Early Relational Health [referred to as Attachment in prior versions of PediaTrac], and Social/Sensory Information Processing) at each of eight sampling periods (NB, 2, 4, 6, 9, 12, 15, and 18 months). Survey questions about demographics, as well as family and perinatal medical characteristics were completed during the NB period, with information on the family environment and infant medical status updated at all subsequent assessments. The time required to complete PediaTrac at each time period is 20–30 min. The focus of this investigation is on the 15 items of the Motor domain at the NB period which required approximately 5 min to complete. The MOT items are administered as a clustered component.

Information describing the original item bank and domain development, expert panel reviews, cognitive interviews with parents, and the pilot validation results of PediaTrac 2.0 have been previously published.[20] Previous analyses of PediaTrac 2.0 indicated that PediaTrac domains had the potential to produce valid and reliable estimates of infant development; however, results also revealed the need for (1) additional and more varied items at each age group, (2) more common items across adjoining assessments, and (3) a study sample that included infants with developmental risks such as preterm birth.

The current version (3.0) of PediaTrac extended the previous version by adding 15- and 18-month assessments that included duplicate items across timepoints to ensure a sufficient sampling of the range of abilities over development and to allow for modeling of trajectories.[21] To provide for more precise estimates of the domains used in PediaTrac, binary choices were replaced by ordinal response options (i.e., 5-point Likert scales). Each of the 15 motor items at the NB period is ordinal (i.e., ordered categorical responses) and caregivers were required to respond to a 5-point Likert scale with response anchors as follows: 1 = never; 2 = rarely; 3 = sometimes; 4 = often; 5 = always. Items were scaled such that higher scores represented more developed motor abilities. The domains assessed in the current version have remained unchanged. The PediaTrac 2.0 item bank (NB through 12-month items) has been significantly revised, with items removed that performed poorly statistically. Finally, to ensure that the latent trait of "development" could be effectively modeled over time and ability level, items were duplicated across two earlier assessments for all but the NB and 2-month periods (which were identical) and at one later assessment. This repetition strategy ensured that items were sufficiently sampling the range of abilities/traits across child development and to allow for a method of yoking consecutive periods in modeling developmental trajectories.

Statistical analyses

IRT modeling was used to identify a latent trait of NB motor ability. Factor structure of the items was further examined with EFA. Descriptive statistics were computed for all demographic variables and IRT-based thetas for the PediaTrac MOT domain.

IRT modeling. IRT modeling within a Bayesian framework was performed using PediaTrac ratings of the 15 items of the NB MOT domain (see Table 2) to assess infant motor development.[29,33] Graded response modeling (GRM) was used to model item parameters.[29] The IRT analysis was performed using SAS (SAS v9.4, SAS Institute, 2016). The items were evaluated using item characteristics and information curves to determine whether to be included or excluded from final IRT models. Item missing data were handled through constructing the observed data log-likelihood from the appropriate posterior distributions.

IRT models the probability of each item response category as a function of item parameters and estimates the infants' latent abilities or traits in a given domain as represented by values of theta (θ).[32] The item parameters modeled are: (1) item discrimination (slope) = a; and (2) item difficulty (location) = b. Item discrimination describes how well an item discriminates between individuals at different levels of the trait. If discrimination is high, then the item can differentiate infants with high and low abilities. Item difficulty indicates the location on the latent trait (x-axis) where the probability is at least 50% of assigning a rating in a given category or higher for that item. Items with higher b values are those endorsed at higher ability levels.

Other information on item characteristics is provided by Item Characteristic Curves (ICC), Item Information Curves (IIC), and the Test Information Curve (TIC). ICC plots across the levels of the latent trait for a

Table 2. Fifteen items of the PediaTrac newborn Motor domain.

1. Can your child lift head while lying on their stomach?
2. When a small toy (or other object) is placed in your child's hand, do they grasp it?
3. Does your child turn their head to left and right?
6. Does your child push their chest up while lying on stomach?
7. Can your child hold an object (e.g., rattle/ring) briefly?
8. Does your child look side to side and up and down?
9. Does your child move both arms and legs equally?
10. Does your child reach for an object/toy?
11. Is your child propping up on their forearms when on stomach?
12. Does your child roll over from stomach to back?
13. Does your child roll over from back to stomach?
14. Does your child usually keep their hands open (not in fist)?
16. Is your child holding their head up without support?
17. Does your child bring fist to mouth?
19. While you gently pull your child up to a sitting position by their hands, does their head flop back?

Items 4, 5, 15, and 18 were eliminated from the newborn Motor domain and IRT analyses given poor distributional properties. Items 4 and 5 had 100% endorsement for scores 4–5 (negatively skewed); items 15 and 18 were endorsed as never nearly 100% of the time.

given item to discriminate individuals with lower vs. higher levels of that trait for each response option. These plots also provide data on item discrimination, such that items with higher discrimination values provide information about the trait in a narrow range.

IIC plots provide information on the extent to which a given item contributes to estimates of the latent trait (theta, θ). Item information is typically highest in the region of the trait near the location of parameter *b*, item difficulty. Certain items may provide more information at low levels of the attribute, while others may provide more information at higher levels. The TIC displays the total information provided by the sum of all of the items along the ability continuum assessed by the domain and permits an estimate of scale reliability based on Rho = (Information/[Information+1]).[34]

Exploratory factor analysis (EFA). Following IRT modeling, EFA was conducted using R to examine the dimensionality and factor structure of the NB Motor domain.[35] A principal axis factor analysis was conducted on the 15 items with an oblique rotation using the following procedures: (1) inspection of the scree plot, (2) inspection of parallel analysis scree plots, which estimates the number of factors by extracting factors until the eigenvalues of the data are less than the corresponding eigenvalues of a same sized, random data set, and (3) retention of factors with eigenvalues >1 (Kaiser's criterion).[36] Item factor loadings were used to assign items to factors. Descriptions of possible motor latent constructs were based on fit indices for all factor solutions based on the following criteria: (1) lowest Bayesian information criterion, (2) Tucker–Lewis Index greater than 0.90 (above 0.85 considered acceptable), and (3) root mean square error of approximation less than 0.05 (below 0.10 considered acceptable).

RESULTS

Item response theory graded response modeling (GRM)

Theta—representing the latent trait of newborn motor ability. Theta values were generated for all 571 participants, reflecting each infant's perceived latent trait of NB motor ability. These values are the primary dependent variable of NB motor ability. The mean (*M*) and standard deviation (SD) theta value for the group was $M = 0.03$; SD = 0.90 and ranged from −3.03 to 3.18.

Reliability and parameter estimates. The 15-item NB Motor domain was highly reliable (0.94), based on the reliability estimate closest to mean theta. IRT parameter estimates for item discrimination (i.e., slope) and item difficulty (i.e., thresholds) are presented in Table 3. Mean theta values and item difficulty threshold values can be thought of as being on a scale similar to a *Z*-score; a distribution centered at zero with a standard deviation metric.[37] Discrimination estimates for the 15 items ranged from 0.43 to 1.93.

The item difficulty parameter threshold is interpreted as the location on the latent trait (*x*-axis) where the probability of assigning a rating in a given category (or higher) for that item is at least 50%. For example, motor item 1, has threshold parameters of: −2.08, −1.36, 0.01, and 1.15. These thresholds constitute cut points for the 5-item Likert categories. The first threshold of −2.08 can be interpreted as a *Z*-score value indicating the probability of a caregiver endorsing this first category for her infant (1 = never) versus scoring in a higher category (2–5), and suggests lower motor skills. Appropriately targeted items for a specific age will have a range of item difficulty parameter threshold values (e.g., −5 to +5). Predominately negative item difficulty parameters across the response categories would reflect relatively easier motor items, while those with predominately positive item difficulty parameters would reflect more difficult items.

Items 1, 2, 6, and 7 appear to best target motor abilities in the NB period when one examines the distribution of beta values (item difficulty) across the four thresholds (*b*1–*b*4). Item 6 appears to be the single best item to measure the latent trait of NB motor ability.

Items 3, 8, 9, and 17 are easier items based on their beta values. These items should be endorsed as at least "sometimes" or a higher response category (e.g., often or always) by most caregivers. In contrast, items 10, 11, 12, 13, and 16 are harder items based on their item difficulty beta values and may better target the motor abilities of infants at 2 months or older. Item 14 has the largest range of item difficulty beta values (−5.24 to 4.65) suggesting that this item may be able to better capture the latent trait of NB motor ability across ability level.

It is important to note that we included both significantly easier (except during the NB period) and more difficult developmental items in each survey period, including items in the NB period assessment that would be appropriate for an infant of 2 or 4 months of age. This was done to ensure that each survey had sufficient range to capture both delayed and precocious development, to ensure appropriate targeting of items from the item bank at each age, and to identify items that could serve as anchors across age ranges in preparation for a computer-adaptive test version. Items 9 and 17 were very rarely endorsed as "never" and had low discrimination as a very low level of ability was required for the caregiver to endorse "always," but these were retained as potential

Table 3. Graded response model IRT estimates for PediaTrac newborn motor domain items

	Item discrimination (SD) (slope)	Item difficulty (SD) (threshold)			
Sensorimotor newborn domain item	***a***	***b1***	***b2***	***b3***	***b4***
1. Can your child lift head while lying on their stomach?	1.448 (0.13)	−2.078 (0.18)	−1.362 (0.12)	0.014 (0.08)	1.153 (0.11)
2. When a small toy (or other object) is placed in your child's hand, do they grasp it?	1.287 (0.12)	−1.246 (0.13)	−0.548 (0.08)	0.420 (0.08)	1.473 (0.14)
3. Does your child turn their head to left and right?	1.226 (0.14)	−5.007 (0.57)	−3.489 (0.34)	−1.794 (0.17)	−0.152 (0.08)
6. Does your child push their chest up while lying on stomach?	1.925 (0.17)	−0.744 (0.07)	−0.240 (0.06)	0.504 (0.06)	1.195 (0.09)
7. Can your child hold an object (e.g. rattle/ring) briefly?	1.631 (0.15)	−0.673 (0.08)	0.067 (0.07)	0.848 (0.10)	1.581(0.13)
8. Does your child look side to side and up and down?	1.366 (0.14)	−2.775 (0.27)	−2.017 (0.20)	−1.132 (0.12)	−0.053 (0.08)
9. Does your child move both arms and legs equally?	0.780 (0.13)	−9.188 (2.06)	−6.600 (1.20)	−3.793 (0.65)	−1.344 (0.23)
10. Does your child reach for an object/toy?	1.468 (0.19)	−0.054 (0.08)	0.823 (0.10)	1.757 (0.18)	2.262 (0.22)
11. Is your child propping up on their forearms when on stomach?	1.531 (0.15)	0.593 (0.08)	1.133 (0.11)	2.035 (0.18)	2.759 (0.26)
12. Does your child roll over from stomach to back?	1.219 (0.22)	2.020 (0.28)	2.833 (0.39)	3.560 (0.50)	4.381 (0.62)
13. Does your child roll over from back to stomach?	1.658 (0.15)	1.679 (0.14)	2.325 (0.21)	3.029 (0.30)	3.750 (0.41)
14. Does your child usually keep their hands open (not in fist)?	0.550 (0.09)	−5.238 (0.98)	−2.593 (0.49)	1.192 (0.27)	4.648 (0.86)
16. Is your child holding their head up without support?	1.072 (0.11)	−0.609 (0.10)	0.601 (0.10)	2.272 (0.21)	3.696 (0.37)
17. Does your child bring fist to mouth?	0.578 (0.09)	−7.823 (1.49)	−5.255 (0.90)	−1.636 (0.29)	1.896 (0.33)
19. While you gently pull your child up to a sitting position by their hands, does their head flop back?	0.428 (0.10)	−4.233 (1.09)	−1.323 (0.43)	2.262 (0.53)	3.850 (0.87)

Items 4, 5, 15, and 18 were eliminated and not included in IRT analyses due to poor distributional properties. Items 4 and 5 had 100% endorsement for scores 1–4 (negatively skewed); items 15 and 18 were endorsed as never nearly 100% of the time.
a alpha, *b* beta.

"red flag" items. Items 14 and 19 have lower item discrimination ($a = 0.55$ and 0.43, respectively), but appear to be sampling traits that are more widely expressed by NB infants with item difficulty thresholds ranging from −4.23 to 3.90. These items may serve as good anchor items across time periods given their ability to effectively sample a wide range of motor trait.

Item characteristic curves (ICC). ICC show where along the continuum of the latent trait, each response option discriminated lower vs. higher levels of the trait. Each ICC plot contains five lines corresponding to the five Likert response categories and is based on the levels of caregiver ratings of infants' motor ability (latent trait) represented by the *x*-axis of the plot (theta). Supplementary Figs. 1–3 illustrate the Motor domain items that appear to best target NB motor ability, items that are relatively easier, and items that are relatively more difficult, based on the location of the distributions and median response options. Supplementary Fig. 4 depicts that Motor domain items 14 and 19 appear best able to sample NB motor ability, although with lower item information.

Item information curves and test information curve for newborn motor domain. IIC were plotted to illustrate how well each item measured the latent trait of motor ability at various levels of the attribute. Items that were well-targeted for the NB Motor domain provided the most information (Supplementary Fig. 5). Item 6 provided the most information of all items, with Information approaching 3.0. Relatively easier items provided much less information (see Supplementary Fig. 6), and while items 9 and 17 provided little information, as noted, these items may serve as important detectors of motor delay. More difficult items provided more information than did easier items (see Supplementary Fig. 7).

A TIC illustrates the sum of the item information functions, reflecting the ability of the survey to sample the overall latent trait, NB motor functioning. Total information for the NB Motor domain was 15 and the reliability of the NB Motor domain was 0.94 (see Fig. 1).

Exploratory factor analyses

Given that the distinctions made between different aspects of early infant motor function do not correspond to gross and fine motor distinctions made in later childhood, and that this instrument is based on caregiver observation, we did not take an a priori approach to examining factor structure. A principal axis factor analysis was conducted on the 15 items with oblique rotation. The Kaiser–Meyer–Olkin (KMO) measure verified the sampling adequacy for the analysis, KMO = 0.8. The KMO statistic ranges from 0 to 1 with values closer to 1 indicating that distinct and reliable factors can be detected. All individual items were 0.67 or higher, above the acceptable range of 0.5.[38] Bartlett's test was significant, χ^2 (105) = 2237.35, $p < 0.001$, which confirmed that the correlation matrix was sufficiently different from the identity matrix to justify EFA. The determinant of the correlation matrix (0.0189535) also exceeded 0.00001, the established cutoff.

Analyses were conducted to obtain scree plots and eigenvalues for each factor in the data. A single factor had an eigenvalue over Kaiser's criterion of 1.0, while the point of inflexion suggested that extracting three factors may be sufficient. We conducted a parallel analysis to explore secondary factors that might distinguish different types of motor skills. We retained four factors as the parallel analysis scree plot suggested a point of inflexion at five factors. Four factors accounted for 43.20% of the variance explained by the extracted factors. Table 4 lists the factor loadings after rotation. The items that cluster on the same factor suggest that factor 1 represents Bilateral Motor Function, factor 2 represents Trunk Control, factor 3 represents Rolling, and factor

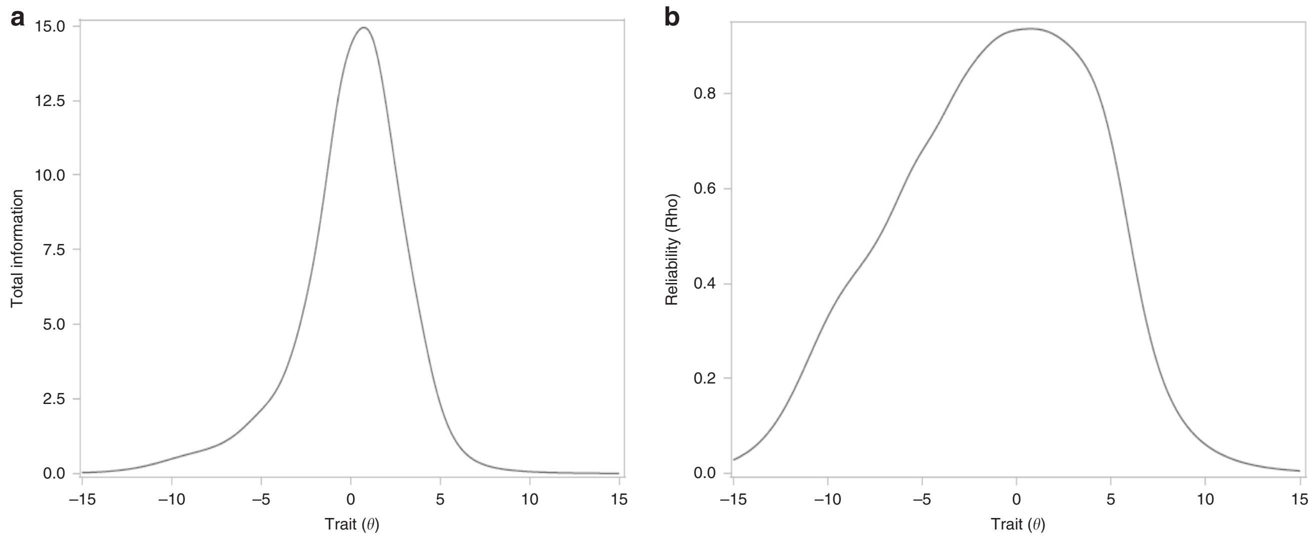

Fig. 1 Test information curve and reliability curve for the PediaTrac newborn Motor domain. a Test Information Curve (TIC) for 15, 5-category items. **b** Reliability Curve (Rho) for 15 5-category items. Mean Theta among 571 respondents = 0.025 and reliability closest to the mean = 0.936.

4 represents Manual Motor Movement. Table 5 reports the fit indices for all four factor solutions.

DISCUSSION

This study represents the first application of IRT GRM to the NB period motor items. The primary aim of this investigation was to model a latent trait of NB motor ability using the PediaTrac ratings of the 15 motor items. Item discrimination and item difficulty could be reliably estimated for all 15 items. In addition, the latent trait, motor ability, could be reliably estimated by theta during the NB period with excellent precision.

GRM identified four items that best targeted NB motor abilities that could be reliably reported by our caregivers. Two of these items were related to general trunk and head control while two addressed manual motor movements. All four of these items provided good information (i.e., reliability). An item that queried about the infant's ability to push their chest up while prone provided the most information of these NB items. These first two items clustered on factor 2, "trunk control," while the last two loaded on factor 4, "manual motor movement." Items that sampled trunk and head control can be thought of as examining the emergence of postural control, the most fundamental motor action.[39] Postural development is the mastery of more complex erect postures over a progressively smaller base of support. These skills are the foundation upon which other motor actions are assembled. However, it is essential to note that there is some variability in the acquisition of motor skills, as milestone charts reflect cultural biases of early researchers and cohorts sampled.[40] In some cultures, infants learn to walk prior to crawling or bypass crawling altogether.[41] The behaviors sampled by the PediaTrac questions that best targeted the NB period are consistent with "mastery over gravity" that precedes top-down in early development, from head to feet.

Inspection of the items that loaded onto factor 1, "bilateral motor function," reveals that they query motor behaviors for which head movement to the left and right, bilateral upper and lower extremity, and full ocular movements (e.g. looking in all directions) are required. The two items related to head and ocular movements provided the most information of these easier items during the NB period. Looking behavior involves the coordination of the infant's body, head, and eyes to bring a location into view, and initial looking is focused on what is directly in front of the infant.[38] Given that NB infants should be able to move their head, arms, and legs equally to the left and right at the NB period, the lack of these abilities are potential developmental "red flags" and could reveal asymmetries (i.e., lack of equalities in movement between the two sides of the body) that warrant follow-up by the pediatrician.

Infants who are born preterm have an increased risk of intraventricular hemorrhage or hypoxic-ischemic events and may exhibit asymmetries suggestive of hemiparesis, motor delay, or a developmental motor disorder.[42] Even mild motor delays in crawling and walking in infancy have been shown to increase the risk for subsequent motor impairments in childhood.[43] However, there is evidence that early motor intervention between 3 and 6 months corrected age in high-risk preterm infants can disrupt this cascade and positively impact subsequent motor and cognitive outcomes.[42] As such, early identification and intervention of even mild motor deviations may be able to disrupt this negative cycle. The current investigation suggests that caregivers can reliably report infant motor behaviors and may be able to play a key collaborative role in identifying possible delays or deviations in development in the NB period, even in infants without known risk.

Two of the motor items, "usually keep hands open," and "when pull to sit, does head flop back?" reflect behaviors that span the widest range of theta, suggesting these items are likely to best capture NB infant motor functioning across the full range of ability. However, these items had poor factor loadings and likely provide qualitatively different information than the remaining items that reflect motor milestones. As research has demonstrated that these motor behaviors are pathognomonic, subsequent modeling may reveal their potential to also serve as specific "red flag" indicators of risk.

Item analyses revealed that five motor items were difficult for NB infants. The items probed the infant's ability to engage in skills such as holding head without support, rolling, propping on forearms, and reaching for objects. These motor behaviors are not typically achieved by NB infants, but they were included in the NB survey during test construction to serve as "yoking" items for the eventual modeling of motor trajectories to the 2- and 4-month periods. Infants who were likely exhibiting these behaviors were those preterm infants who were up to 8 weeks of age (chronological) when caregivers completed the NB survey. Despite these being difficult items for NB infants, all five items provided good information. Of these five items, two items loaded onto factor 2, trunk control, and one item loaded onto factor 4,

Table 4. Summary of exploratory factor analysis results for the PediaTrac newborn Motor domain ($N = 571$).

	Bilateral motor function	Trunk control	Rolling	Manual motor movt.		
Item	Factor 1	Factor 2	Factor 3	Factor 4	Com.	Content
mot 08	**0.75**	0.01	0.03	0.05	0.61	Look side-side up-down
mot 09	**0.73**	−0.06	−0.02	−0.07	0.48	Move both arms-legs equal
mot 03	**0.61**	0.09	−0.02	0.03	0.43	Turn head left-right
mot 17	0.23	0.12	0.08	−0.01	0.10	Bring fist to mouth
mot 14	0.19	0.10	0.11	0.05	0.11	Usually keep their hands open
mot 01	0.02	**0.70**	−0.10	0.02	0.47	Lift head while on stomach
mot 16	−0.03	**0.63**	0.07	−0.07	0.38	Holding head without support
mot 06	0.10	**0.58**	0.01	0.12	0.48	Push up while on stomach
mot 11	−0.03	**0.49**	0.24	0.04	0.40	Prop-up on forearms while on stomach
mot 19	−0.11	−0.20	−0.05	0.06	0.07	When pull to sit, does head flop back
mot 13	0.02	−0.03	**0.85**	0.04	0.72	Roll over back to stomach
mot 12	−0.01	0.03	**0.80**	−0.04	0.64	Roll over stomach to back
mot 07	−0.01	−0.02	−0.01	**0.95**	0.87	Hold object briefly
mot 02	0.03	0.10	0.002	**0.53**	0.37	Grasp small toy
mot 10	0.06	0.15	0.15	**0.41**	0.35	Reach for object

mot motor, *Com.* Communality, *Movt.* movement.
Factor loadings over 0.40 appear in bold.

Table 5. Fit indices for all factor solutions for the PediaTrac newborn Motor domain items.

Factors	Cumulative	χ^2	*p*	BIC	TLI	RMSEA	Lower	Upper
1	23.33	909.32	0	338.06	0.55	0.13	0.12	0.13
2	31.70	556.68	0	74.27	0.69	0.11	0.10	0.11
3	37.92	288.89	0	−110.99	0.82	0.08	0.08	0.09
4	43.20	128.02	0	−195.70	0.93	0.05	0.04	0.06

Cumulative is the cumulative percent of variance explained by the factor solution. Lower and upper are the limits of a 90% confidence interval for RMSEA.
BIC Bayesian information criterion, *TLI* Tucker–Lewis index, *RMSEA* root mean square error of approximation.

reflecting manual motor abilities. The ability to maintain head position and control while being held by a caregiver allows the infant to maintain gaze with them and visually explore their environment.[44] The remaining two items constituted factor 3, rolling. These items would typically reflect more advanced postural control in typically developing infants, but may also be pathological if poorly executed, again potentially serving as "red flags." Rolling becomes one of the first forms of mobility and evidence of the infant's ongoing ability to overcome gravity. Locomotion, including rolling, is not hardwired, is achieved in unique ways, and improves over time with practice.[45]

Given that PediaTrac has been developed with IRT methods, motor ability can be reliably measured and interpreted at both the item and domain level. This will provide clinicians with a metric of the infant's level of ability for each motor behavior or an aggregate index of their motor ability at the time period measured. The PediaTrac Motor domain score can be used independently in the primary care setting during the NB period and would require only minutes to complete. Psychometric studies of the PediaTrac Motor domain for additional time periods that correspond to well child visits are currently underway, which will allow for an eventual examination of motor trajectories over early development. Plans for a streamlined computer-adaptive version of PediaTrac will also allow primary care providers to examine several domains of early development concurrently and with less burden.

This is the first known caregiver instrument that can assess motor ability in the NB period with high reliability in typically developing and at-risk infants. However, it is not without limitations. Differential reporting patterns have been described in caregiver reports,[27] especially in socio-demographically diverse populations. Unlike other online diagnostic and healthcare management systems, PediaTrac has the potential to model the effect of these biases on ratings.[26,27] Embedded response style (i.e., bias) indices are being developed to examine the degree to which the caregiver has a positive or negative response style as well as an atypical response style (e.g., not attending).[21] In addition, further validation of the motor domain at each timepoint will be necessary. In order to examine the predictive validity of our latent motor construct, direct neurodevelopmental assessments are occurring at 24 months of age in a subset of the participants as part of the larger study.

In conclusion, this is the first study that reveals that caregivers can reliably report NB motor functioning. The PediaTrac NB Motor domain could advance knowledge about how motor problems emerge during infancy, identify factors related to their emergence, and examine ways in which trajectories of motor development in at-risk infants differ from those of typically developing infants. With this knowledge, risk for disorders that involve early motor delays, such as CP, can be identified sooner and intervention initiated earlier in infancy.

DATA AVAILABILITY

The datasets generated during and/or analyzed during the current study are not publicly available yet as this is a longitudinal investigation that is still underway. Upon completion of the longitudinal investigation, the datasets generated during this study will be deposited into the National Database for Autism Research (NDAR).

REFERENCES

1. Grantham-McGregor, S. et al. Developmental potential in the first 5 years for children in developing countries. *Lancet* **369**, 60–70 (2007).
2. Almas, A., Degnan, K. A., Nelson, C. A., Zeanah, C. H. & Fox, N. A. IQ at age 12 following a history of institutional care: findings from the Bucharest Early Intervention Project. *Dev. Psychol.* **52**, 1858–1866 (2016).
3. Ferrari, F. et al. Motor and postural patterns concomitant with general movements are associated with cerebral palsy at term and fidgety age in preterm infants. *J. Clin. Med.* **8**, 1189 (2019).
4. Ben-Itzchak, E. & Zachor, D. A. Toddlers to teenagers: long-term follow-up study of outcomes in autism spectrum disorder. *Autism* **24**, 41–50 (2020).
5. Posar, A. & Visconti, P. Long-term outcome of autism spectrum disorder. *Turk. Pediatr. Ars.* **54**, 207–212 (2019).
6. Dennis, M. et al. Functional plasticity in childhood brain disorders: when, what, how, and whom to assess Maureen. *Neuropsychol. Rev.* **24**, 389–408 (2014).
7. Scattone, D., Raggio, D. J. & May, W. Comparison of The Vineland Adaptive Behavior Scales, second edition, and the Bayley scales of infant and toddler development, third edition. *Psychol. Rep.* **109**, 626–634 (2011).
8. Beker, L. Principles of growth assessment. *Pediatr. Rev*. **27**, 196–197 (2006).
9. Zubler, J. M., et al. Evidence-informed milestones for developmental surveillance tools. *Pediatrics*. **149**, e2021052138. https://doi.org/10.1542/peds.2021-052138 (2022).
10. Allison, K. M., Cordella, C., Iuzzini-Seigel, J. & Green, J. R. Differential diagnosis of apraxia of speech in children and adults: a scoping review. *J. Speech Lang. Hear Res.* **63**, 2952–2994 (2020).
11. Brothers, K. B., Glascoe, F. P. & Robertshaw, N. S. PEDS: developmental milestones-an accurate brief tool for surveillance and screening. *Clin. Pediatr. (Philos.)*. **47**, 271–279 (2008).
12. Bellman, M., Byrne, O. & Sege, R. Developmental assessment of children. *BMJ* **345**, 31–35 (2013).
13. Shevell, M. et al. Evidence Report: genetic and metabolic testing on children with global developmental delay: report of the quality standards Subcommittee of the American Academy of Neurology and the Practice Committee of the Child Neurology Society. *Neurology* **60**, 367–380 (2003).
14. Allotey, J. et al. Cognitive, motor, behavioural and academic performances of children born preterm: a meta-analysis and systematic review involving 64 061 children. *BJOG Int J. Obstet. Gynaecol.* **125**, 16–25 (2018).
15. Novak, I. et al. Early, accurate diagnosis and early intervention in cerebral palsy: advances in diagnosis and treatment. *JAMA Pediatr.* **171**, 897–907 (2017).
16. Romeo, D. M. et al. Early neurological assessment and long-term neuromotor outcomes in late preterm infants: a critical review. *Medicina (Kaunas)* **56**, 1–13 (2020).
17. Bayley, N. & Aylward, G. P. *Bayley Scales of Infant and Toddler Development* 4th edn (NCS Pearson, Inc., Bloomington, MN, 2019).
18. Mullen, E. M. *Mullen Scales of Early Learning* (Circle Pines, MN, American Guidance Service Inc. National Research Council, 1995).
19. Macy, M. L., Huetteman, P. & Kan. K. Changes in primary care visits in the 24 weeks after COVID-19 stay-at-home orders relative to the comparable time period in 2019 in Metropolitan Chicago and Northern Illinois. *J. Prim. Care Community Heal.* **11**, 2150132720969557 (2020).
20. Lajiness-O'Neill, R., et al. Development and validation of PediaTrac™: a web-based tool to track developing infants. *Infant Behav. Dev*. **50**, 224–237. https://doi.org/10.1016/j.infbeh.2018.01.008 (2018).
21. Lajiness-O'Neill, R. et al. PediaTrac V.3.0 protocol: a prospective, longitudinal study of the development and validation of a web-based tool to measure and track infant and toddler development from birth through 18 months. *BMJ Open.* **11**, e050488 (2021).
22. Squires, J., Bricker, D. & Twombly, E. Ages & Stages Questionnaires: a parent-completed child monitoring system-3 (Baltimore, MD, Brooks Publishing Company, 2009).
23. da Silva, M. A., de Mendonça Filho, E. J., Mônego, B. G. & Bandeira, D. R. Instruments for multidimensional assessment of child development: a systematic review. *Early Child Dev. Care.* **190**, 1257–1271 (2020).
24. Sparrow, S., Cicchetti, D. V. & Saulnier C. A. *Vineland Adaptive Behavior Scale* 3rd edn (VABS-3). (San Antonio, TX, Pearson, 2016).
25. Voress, J. K. & Maddox, T. *Developmental Assessment of Young Children* 2nd edn (DAYC-2). (Austin, TX, PRO-ED, 2013).
26. Child Health & Development Interacteractive System (CHADIS). http://www.chadis.com/.
27. Cella, D. et al. The Patient-Reported Outcomes Measurement Information System (PROMIS). *Med. Care.* **45**(5 Suppl 1), S3–S11 (2007).
28. Embretson, S. E. & Reise, S. P. *Item Response Theory for Psychologists* (Erlbaum, Mahwah, NJ, 2000).
29. Samejima, F. Graded response model. In *Handbook of Modern Item Response Theory* (eds van der Linden, W. J. & Hambleton, R. K.) (Springer, New York, NY, 1997). https://doi.org/10.1007/978-1-4757-2691-6_5
30. Kamata, A. & Bauer, D. J. A note on the relation between factor analytic and item response theory models. *Struct. Equ. Model.* **15**, 136–153 (2008).
31. Thissen, D. & Orlando, M. Item response theory for items scored in two categories. In *Test Scoring* 1st edn (eds Thissen, D., Nelson, L., Rosa, K. & McLeod, L. D.) (Routledge, New York, 2001). https://doi.org/10.4324/9781410604729
32. Hambleton, R. K., Swaminathan, H. & Jane Rogers, H. *Fundamentals of Item Response Theory* (Sage Publications, Inc., Newbury Park, CA, 1991).
33. de Ayala, R. J. The theory and practice of item response theory. *Ref. Res. B News*. **24**, 209–236 (2009).
34. Nguyen, T. H., Han, H.-R., Kim, M. T. & Chan, K. S. An introduction to item response theory for patient-reported outcome measurement. *Patient* **7**, 23–35 (2014).
35. R Core Team. *R: A Language and Environment for Statistical Computing* (Vienna, Austria, 2020).
36. Kaiser, H. F. The application of electronic computers to factor analysis. *Educ. Psychol. Meas.* **20**, 141–151 (1960).
37. Depaoli, S., Tiemensma, J. & Felt, J. M. Assessment of health surveys: fitting a multidimensional graded response model. *Psychol. Heal Med.* **23**, 13–31 (2018).
38. Gibson, E. Exploratory behavior in the development of perceiving, acting, and the acquiring of knowledge. *Annu Rev. Psychol.* **39**, 1–41 (1988).
39. Adolph, K. E. & Franchak, J. M. The development of motor behavior. *Wiley Interdiscip. Rev. Cogn. Sci.* **8**, 1–30 (2017).
40. Adolph, K. E., Karasik, L. B. & Tamis-Lemonda, C. S. Motor skill. In: *Handbook of Cultural Developmental Science* (ed. Bornstein, M. H.) 61–88 (Psychology Press, 2010).
41. Hopkins, B. & Westra, T. Motor development, maternal expectations, and the role of handling. *Infant Behav. Dev.* **13**, 117–122 (1990).
42. Blauw-Hospers, C. H., de Graaf-Peters, V. B., Dirks, T., Bos, A. F. & Hadders-Algra, M. Does early intervention in infants at high risk for a developmental motor disorder improve motor and cognitive development. *Neurosci. Biobehav Rev.* **31**, 1201–1212 (2007).
43. Hua, J. et al. Early motor milestones in infancy and later motor impairments: a population-based data linkage study. *Front Psychiatry* **13**, 1–11 (2022).
44. Fogel, A., Messinger, K., Dickson, L. & Hsu, H. Posture and gaze in early mother-infant communication: synchronization of developmental trajectories. *Dev. Sci.* **2**, 325–332 (2001).
45. Rachwani, J. et al. Segmental trunk control acquisition and reaching in typically developing infants. *Exp. brain Res.* **228**, 131–139 (2013).

ACKNOWLEDGEMENTS

The investigators express their sincere gratitude to caregivers who participated in the PediaTrac project. We would like to acknowledge the Corner Health Center for their support in helping recruit those with the greatest need.

AUTHOR CONTRIBUTIONS

Conception and design of the study: R.L.-O., T.R., S.W., P.B., A.H.-B., H.G.T., J.C.G.L., J.B., and A.L. Acquisition, analysis, and interpretation of data: T.R., P.B., A.D.S., J.C.G.L., R.L.-O., S.W., A.H.-B., and H.G.T.. Drafting and revising the manuscript (original): R.L.-O., T.R., P.B., S.W., A.H.-B., H.G.T., A.S., and J.C.G.L. Final approval: R.L.-O., T.R., P.B., A.H.-B., H.G.T., A.S., J.B., A.L., J.C.G.L., and S.W.

FUNDING

This study was supported by the Eunice Kennedy Shriver National Institute of Child Health & Human Development of the National Institutes of Health under Award Number (R01HD095957). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. Efforts of S.W. were also supported in part by the Mildred E. Swanson Foundation.

COMPETING INTERESTS

The authors declare no competing interests.

ETHICS APPROVAL AND CONSENT TO PARTICIPATE

Participant consent was required, with the study approved by the University of Michigan IRB of Record (HUM00151584).

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41390-022-02312-4.

Correspondence and requests for materials should be addressed to Renee Lajiness-O'Neill.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

PEDIATRAC PROJECT CONSORTIUM

Casey Swick[1], Samantha Goldstein[1], Michelle Lobermeier[1], Amanda Hicks[1], Jennifer Cano[1], Shannon Franz[2], Najae Dixon[2], Kirsten Oard[2], Lesa Dieter[4], Jazmine Kirkland[4], Yanisa Robbins[4] and Emily Gorjanc[4]