

# Tummy Time Tracking: Examining Agreement Between Parent Recall and Direct Observation in Infants

Ketaki Inamdar, Stacey C. Dusing, Leroy Thacker, Peter E. Pidcoe, Sheryl Finucane, Virginia Chu

**Importance:** Parent recall is the primary method for measuring positioning practices such as tummy time in infants. Concerns regarding the accuracy of parent recall have been raised in the literature. To date, no study has examined the agreement of tummy time recall measures with gold-standard methods.

**Objective:** To assess the agreement between parental recall versus direct observation of tummy time in infants, and to explore the impact of prematurity on this relationship.

**Design:** Cross-sectional observational study, spanning 1 yr.

**Setting:** Participants' homes

**Participants:** Thirty-two infant–parent dyads (19 full-term, 13 preterm), with infants ages 3 to 6 mo and caregivers ages older than 18 yr.

**Outcome and Measures:** Home-recorded videos of infant play across 3 days were used as a proxy for direct observation of tummy time and compared with a 12-item parent recall survey.

**Results:** Parent recall had a significant moderate correlation ( $\rho = .54, p = .002$ ) with direct observation in full-term infants but was not correlated ( $\rho = .23$ ) with direct observation in preterm infants. On average, parents of preterm infants overestimated tummy time by 2.5 times per day compared with direct observation.

**Conclusions and Relevance:** For full-term infants, parent recall measures of tummy time exhibit an acceptable level of agreement with direct observation and can be reliably used over shorter periods. Parents of preterm infants may display a bias in recalling tummy time, leading to overestimations. To accurately assess tummy time in this population, a combination of subjective and objective measures should be explored.

**Plain-Language Summary:** Tummy time is an essential movement experience for infants, especially for preterm infants, who are at a higher risk for motor delays. The most common way to track tummy time is through parent reports, or recall, versus a practitioner directly observing tummy time in the home. Despite the widespread use of parent recall to track tummy time, no study has examined the accuracy of parent recall versus direct observation in the home. Accurately assessing tummy time is crucial for improving and supporting health outcomes for infants. This study found that prematurity may affect the accuracy of parent recall for assessing tummy time in young infants. The authors discuss the implications of this finding and provide suggestions to guide the selection of appropriate methods to measure tummy time in clinical practice and research studies.

Inamdar, K., Dusing, S. C., Thacker, L., Pidcoe, P. E., Finucane, S., & Chu, V. (2024). Tummy time tracking: Examining agreement between parent recall and direct observation in infants. *American Journal of Occupational Therapy*, 78, 7804185020. <https://doi.org/10.5014/ajot.2024.050521>

**M**ovement experiences in early childhood (0–5 yr) have numerous health benefits, including improved global development, reduced adiposity, and enhanced bone and cardiometabolic health (Carson et al., 2017). Play, the most essential occupation of a child, functions as a medium through which these pivotal movement experiences come to fruition. The

World Health Organization's 24-hr movement guidelines (Sommer et al., 2021) and the American Academy of Pediatrics' (AAP's) "Back to Sleep, Tummy to Play" campaign (American Academy of Pediatrics, 2016) advocate for at least 30 min of daily active prone play, or *tummy time*, as the primary play experience for young infants who are not yet mobile.

Tummy time is essential for infants, particularly for those born preterm (<37 wk of gestation), as they are at a heightened risk of motor delays. Engaging in tummy time offers several opportunities for infants to develop antigravity strength and stability and, thus, lays the groundwork for foundational gross motor skills (Russell et al., 2009). As infants learn to push themselves up using their arms during tummy time, they enhance their upper body strength and coordination which is pivotal for later fine motor development (Lee, 2015). Propping up on the tummy also helps infants to practice visual tracking, a skill that is crucial for developing eye coordination and depth perception (Senju et al., 2018). Research shows that the dosage of tummy time (i.e., the amount of time an infant spends on their tummy) is crucial for achieving optimal outcomes. Full-term infants who engage in tummy time for 30 to 90 min/day achieve motor milestones earlier than those who engage for less than 30 min (Dudek-Shriber & Zelazny, 2007; Russell et al., 2009). Similarly, preterm infants at an adjusted age of 8 mo, who spend more than 40 min/day on their tummy exhibit superior gross motor skills, compared with those with less tummy time (Bartlett & Fanning, 2003). Engaging in tummy time for more than 5 min/day is also recognized as a protective factor against plagiocephaly in infants. Infants with severe deformational plagiocephaly who meet this tummy time threshold show a 46% reduction in the severity of plagiocephaly at age 6 mo (van Vlimmeren et al., 2008). Collectively, these findings highlight the importance of tummy time dosage in achieving targeted health outcomes, and they underscore the need for accurate measurement of tummy time at home.

The most common method used to track tummy time duration at home is through parent reports, which are typically obtained through retrospective recall ranging from 24-hr recall to a month (Dudek-Shriber & Zelazny, 2007). Parent recall measures are suitable for capturing the sporadic distribution of tummy time in various contexts, making them cost-effective and time efficient for use in longitudinal and population-based studies (Bennetts et al., 2016). However, concerns regarding the validity of these recalls have been raised, because parent-reported dosages often do not align with gold-standard methods such as direct observation. For example, Kippe et al. (2022) found that only 5% of parents of children ages 4 to 6 yr accurately estimated their children's physical activity levels, with the majority overestimating their child's activity. Parent recall measures are also susceptible to subjectivity (van Zyl et al., 2016) and social desirability bias (Lundahl et al., 2014), and they are moderated by personal and social factors such as parental mental health (Harvey et al., 2013), education (Reese & Read, 2000), and socioeconomic status (Bornstein et al., 2020).

Despite the widespread use of parent recall measures for tummy time, no study has examined their

agreement with direct observation in home settings. This research gap is significant, because inaccurate measurement of tummy time could result in the underestimation or overestimation of adherence patterns in infants. Notably, many tummy time intervention studies rely on parent recall as an outcome measure (Palmer et al., 2019; Uzark et al., 2021), making it imperative to examine its suitability as an outcome measure in clinical practice and research studies. The first objective of this study was to quantify the agreement between a 3-day tummy time parent recall and directly observed tummy time in home settings. The second objective was to compare the accuracy of tummy time parent recall between the parents of full-term and preterm infants.

## Method

### Participants

Thirty-two infants—19 full-term (age,  $M = 5.35$  mo,  $SD = 1.17$ ) and 13 preterm (gestational age,  $M = 31.62$  wk,  $SD = 3.66$ ; adjusted age,  $M = 4.60$  mo,  $SD = 1.02$ )—and their parents participated in this study. They were recruited through convenience sampling from hospital and community settings using flyers, social media advertisements, and mailed letters. All infants were a part of a broader investigation aimed at validating different measures of tummy time. Thus, the determination of a sample size was guided by the intra-class coefficient (ICC) values gleaned from established studies (Hewitt et al., 2019; Ricardo et al., 2018). Although our original intention was to attain a balanced representation of both full-term and preterm infants, recruitment during the COVID-19 pandemic significantly curtailed our capacity to obtain a larger cohort of preterm infants. Descriptive statistics for the infants and their caregivers are presented in Table 1.

Infants ages 3 to 6 mo (adjusted age for preterm infants) were included if their caregivers were English-speaking adults ages 18 or older and gave consent for video and audio recording. Infants who were intolerant to tummy time (crying for over 30 s when placed on the tummy for a 1-min period) or who had medical conditions that prevented lying on their tummy were excluded. Written parental consent was obtained, and the study was conducted in accordance with the Declaration of Helsinki and approved by the Virginia Commonwealth University Institutional Review Board (IRB No. HM20020592).

### Measures

#### *Direct Observation (Gold Standard)*

A tripod-mounted camcorder (Panasonic HC-770) was used as a proxy for direct observation of tummy time at home. This camcorder can operate while being charged, enabling continuous recording for extended durations. The video data captured for each infant were analyzed frame by frame using the Datavyu

**Table 1. Participant Characteristics**

Characteristic	n (%)		
	All Infants (n = 32)	Full-Term Infants (n = 19)	Preterm Infants (n = 13)
Infant age, mo, <i>M (SD)</i> <sup>a</sup>	5.04 (1.16)	5.35 (1.17)	4.60 (1.02)
Infant birth weight, lb, <i>M (SD)</i>	5.78 (1.92)	6.9 (0.85)	3.96 (1.73)
Gestational age, wk			
>37	19 (59.4)	19 (100)	0 (0)
32–37	6 (18.7)	0 (0)	6 (46.2)
28–32	5 (15.6)	0 (0)	5 (38.4)
<28	2 (6.3)	0 (0)	2 (15.4)
Infant ethnicity			
Hispanic	0 (0)	0 (0)	0 (0)
Not Hispanic	30 (93.7)	19 (100)	11 (84.6)
Not reported	2 (6.3)	0 (0)	2 (15.4)
Infant race			
White	19 (59.4)	15 (78.9)	4 (30.8)
Black	4 (12.5)	0 (0)	4 (30.8)
Asian	1 (3.1)	0 (0)	1 (7.7)
Multiracial	5 (15.6)	4 (21.1)	1 (7.7)
Not reported	3 (9.4)	0 (0)	3 (23.0)
Parent age, yr			
18–35	25 (78.1)	16 (84.2)	9 (69.2)
36–55	4 (12.5)	1 (5.3)	3 (23.1)
≥56	1 (3.1)	1 (5.3)	0 (0)
Not reported	2 (6.3)	1 (5.3)	1 (7.7)
Parent ethnicity			
Hispanic	0 (0)	0 (0)	0 (0)
Not Hispanic	30 (93.7)	19 (100)	11 (84.6)
Not reported	2 (6.3)	0 (0)	2 (15.4)
Parent race			
White	22 (68.7)	17 (89.4)	5 (38.4)
Black	4 (12.5)	0 (0)	4 (30.8)
Asian	2 (6.3)	1 (5.3)	1 (7.7)
Multiracial	1 (3.1)	1 (5.3)	0 (0)
Not reported	3 (9.4)	0 (0)	3 (23.1)
Parent education			
High school or diploma, no degree	6 (18.8)	3 (15.8)	3 (23.1)
Associate's degree	3 (9.4)	1 (5.2)	2 (15.3)
Bachelor's or master's degree	15 (46.8)	9 (47.4)	6 (46.2)
Doctorate or professional degree	7 (21.9)	6 (31.6)	1 (7.7)
Not reported	1 (3.1)	0 (0)	1 (7.7)

<sup>a</sup>Adjusted age used for preterm infants.

software (<https://www.datavyu.org>) to identify instances of tummy time, regardless of whether the infant remained stationary or mobile, throughout the 3-day period. The following positions were categorized as

tummy time: prone position with the head in midline or turned to the side, prone on forearms, and prone on hands. Considering the age of our participants, variations of the 4-point position—including initial

attempts to push up on hands and knees, or immature quadruped (Piper & Darrah, 1994)—were also included in the tummy time category. Figure 1 illustrates the commonly observed tummy time variations among infants in our study. Video coders included a Doctor of Occupational Therapy graduate student and an undergraduate pre-health student with a developmental psychology interest who underwent 30-day training. Training involved identifying distinct play positions in infants, as well as recognizing variations in these positions, using test videos and a coding manual. Their coding accuracy was evaluated against an experienced physical therapist, using seven to 10 videos for testing. Feedback was continually provided until a desired reliability benchmark of 90% was achieved (interrater ICC = .91; intrarater ICC = .95).

### *Parent Recall for Tummy Time*

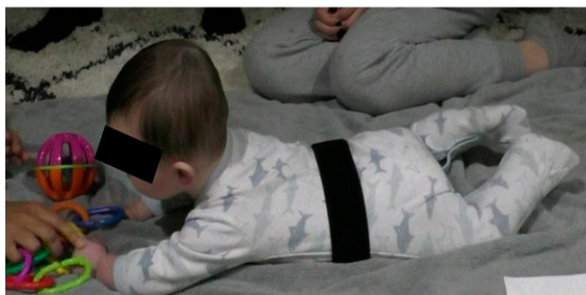
A 12-item parent recall survey (available online as Supplemental Material with this article at <https://research.aota.org/ajot>) was used to collect data on tummy time duration and was administered on the last day of the study through REDCap (Harris et al., 2019). In the first section, parents were presented with images of eight age-appropriate play positions and asked to indicate the amount of time their baby spent in each position during the recorded play sessions on that specific day. The second section focused specifically on tummy time. Parents were requested to recall and report the average daily tummy time their baby engaged in during the recorded play sessions over a 3-day period. They were asked to exclude any tummy time that was not captured on camera. Parents reported

**Figure 1. Variations in tummy time positions observed among infants in our study.**

**A) Prone prop**



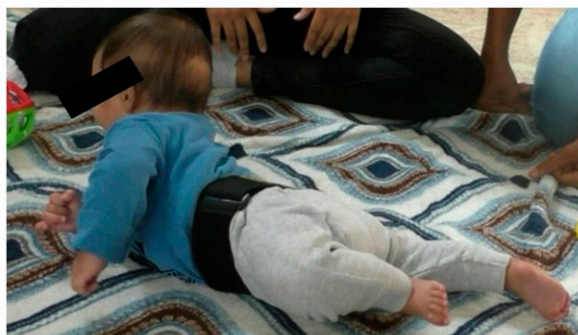
**B) Forearm support**



**C) Prone on extended arms**



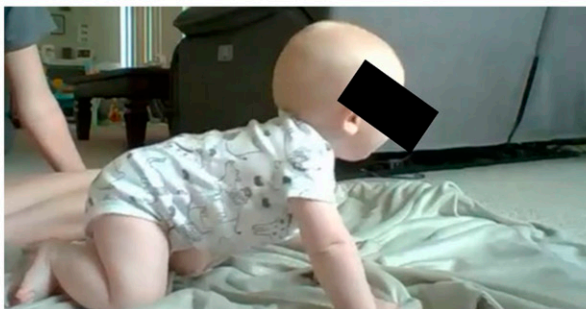
**D) Transition from prone**



**E) Baby plank**



**F) Immature 4-point**





the durations in hours and minutes. Although the survey encompassed information about all play positions to reduce bias related to tummy time, only the tummy time data were used for analysis. The first section of the survey was adapted from a previous study that demonstrated acceptable test–retest reliability (Hesketh et al., 2015); for the ICC ( $M = .57$ , range = .20–.92). The second section drew from established parent recall measures used in prior research on assessing tummy time in infants (Bartlett & Fanning, 2003; Dudek-Shriber & Zelazny, 2007; Russell et al., 2009). Before administration, the face validity of the survey was assessed with the guidelines by Patel and Desai (2020) by a panel that comprised three experienced pediatric physical therapists, an occupational therapist, and a parent.

### Study Procedure

This cross-sectional study was conducted in participants' homes over 3 days during the COVID-19 pandemic (April 2021–March 2022). We implemented a minimal-contact validation protocol to comply with physical distancing regulations. A study package, including a camcorder and an illustrated study manual, was either delivered or mailed to the participants on Day 1. Parents received orientation on camera placement through in-person visits or virtual video conferencing.

During play sessions, parents set up the camera in their infant's play area and recorded the sessions, replicating their daily routines. The camera was turned off during activities such as clothing changes, diaper changes, bathing time, and sleeping to respect privacy. At the end of Day 3, parents completed the parent recall survey. The study package was collected, and the video data were downloaded and marked to identify instances of tummy time. Tummy time duration was calculated for each day and averaged over the 3 days. This average duration was compared with parent recall for agreement. Additionally, 20% of the videos were independently coded, achieving excellent interrater (ICC = .93) and intrarater (ICC = .97) reliability.

### Statistical Analyses

We conducted the statistical analyses using the JMP<sup>®</sup> Pro (Version 15.1.0), and the level of significance was set at .05 or less. Descriptive analyses were reported for sample demographics. Agreement between the two measurement approaches (parent recall and direct observation) was assessed in three steps. First, we used Spearman's rank correlation analyses to evaluate the associations between the approaches, and the strength of correlation coefficients ( $\rho$ s) was interpreted as weak (.1–.4), moderate (.4–.7), or strong (.7–1.0; Akoglu, 2018). Second, Bland–Altman plots were used to visualize and report the systematic bias between the approaches with 95% limits of agreement (Bland & Altman, 1999). Last, linear mixed-model (LMM)

analyses determined the absolute mean differences between the approaches and evaluated whether birth status (full term or preterm) moderated the differences. Least-square mean estimates ( $\beta$  and  $SE$ ), 95% confidence intervals (95% CIs), and fixed-effects results are reported. If the interaction term was significant, we used Student's  $t$  tests for pairwise comparisons. In addition to the  $p$  value of mean differences, percentage accuracy between the two approaches was calculated using the formula percentage accuracy =  $100 - [(\text{mean absolute difference})/\text{direct observation} \times 100]$ . Ideally, two methods are required to have a percentage accuracy of 90% or higher for them to be used interchangeably (Lorne et al., 2018). Given the inherent subjectivity error of parent reports, the 3-day recall burden, and previously reported lower accuracy in children (Kippe et al., 2022), we operationally defined a percentage accuracy of 80% or higher as clinically acceptable.

### Results

Parent recall data were missing for 3 participants: Two parents forgot to complete the form, and one could not recall the tummy time duration. Thus, the final sample consisted of 29 infants ( $n = 18$  full term and 11 preterm). For full-term infants, most parent recall surveys were completed by the infants' mother ( $n = 17$ ; 94%), and one was completed by a grandparent ( $n = 1$ ; 6%). For preterm infants, two surveys were completed by the infants' father ( $n = 2$ ; 18%), and the rest were completed by mothers ( $n = 9$ ; 82%).

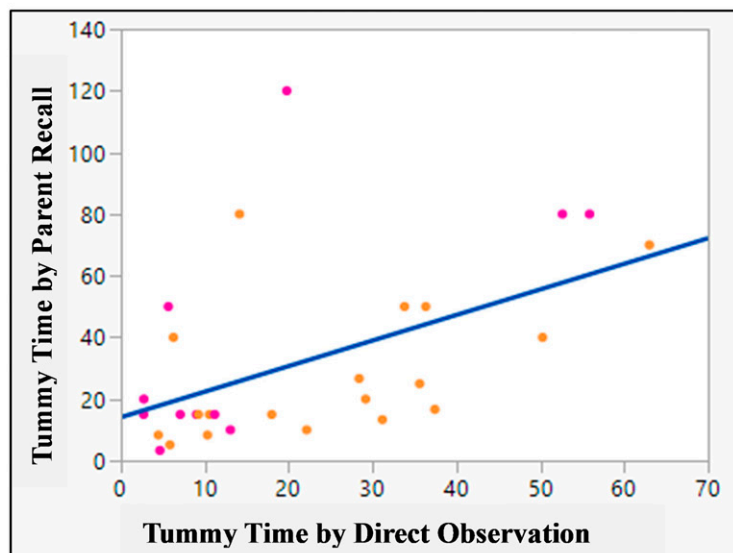
On the basis of the recorded play sessions, the average daily playtime for all infants was 54.59 min ( $SD = 30.14$ ). Full-term infants had a higher average playtime of 65.87 min/day ( $SD = 29.05$ ), compared with preterm infants, who had an average playtime of 38.12 min/day ( $SD = 24.21$ ). Tummy time durations varied among all infants, with a median of 13.48 min/day (interquartile range [IQR] = 6.5, 33.1). When birth status was considered, full-term infants had a higher median tummy time of 23.28 min/day (IQR = 10.3, 35.6) compared with preterm infants, who had a median tummy time of 8.9 min/day (IQR = 4.4, 16.4).

The scatterplot for correlations between parent recall and direct observation of tummy time is presented in Figure 2A. Among all infants ( $n = 29$ ), parent recall had a significant ( $p = .002$ ) and moderate positive ( $\rho = .54$ ) correlation with direct observation. When focusing on full-term infants ( $n = 18$ ), we found that a similar pattern emerged with a significant ( $p = .01$ ) moderate positive ( $\rho = .58$ ) correlation between parent recall and direct observation. Among preterm infants ( $n = 11$ ), however, the correlation between parent recall and direct observation, although moderate ( $\rho = .49$ ), did not reach statistical significance ( $p = .23$ ).

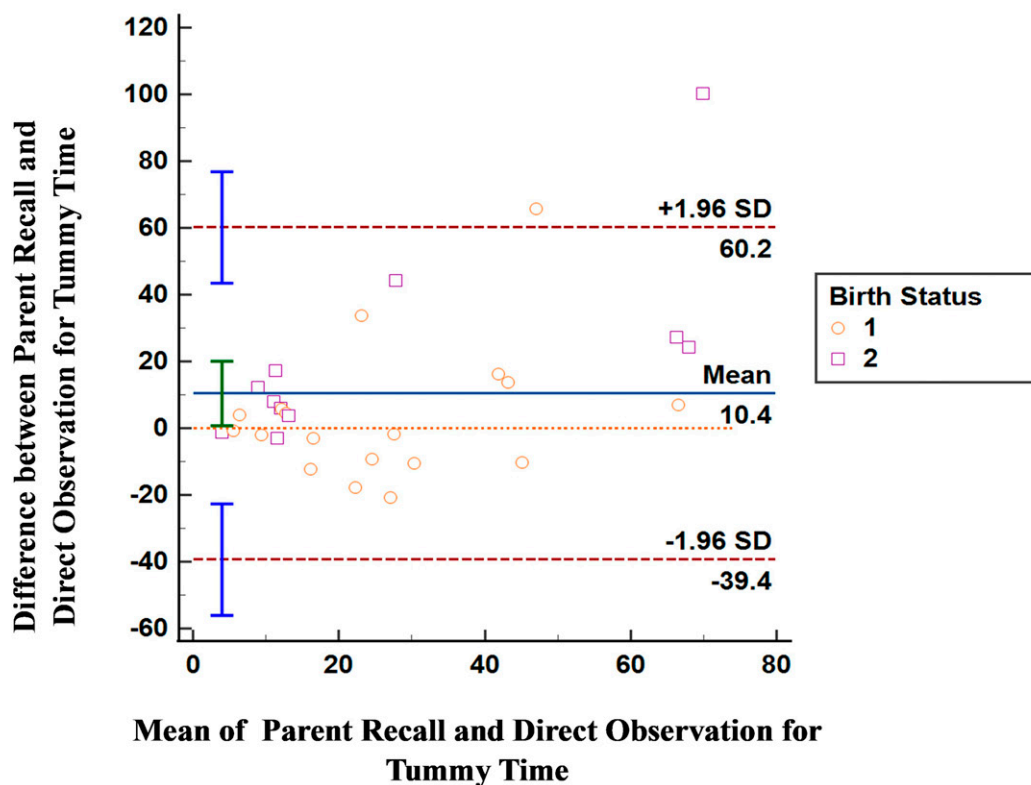
The Bland–Altman plot analysis (Figure 2B) revealed a significant ( $p = .03$ ) overestimation of tummy time duration by parents compared with direct

Figure 2. Scatterplot and Bland–Altman plot of parent recall for tummy time plotted against direct observation for all infants ( $n = 29$ ). (A) Duration of parent-reported tummy time ( $y$ -axis) versus directly observed tummy time ( $x$ -axis) in minutes. The identity line is indicated in blue ( $\rho = .54, p = .002$ ), full-term infants are indicated in orange, and preterm infants are indicated in pink. (B) Difference between the two approaches ( $y$ -axis) versus mean of the two approaches ( $x$ -axis).

**A) Scatter plot (n=29)**



**B) Bland Altman plot (n=29)**



*Note.* Solid blue line = systematic bias between the two approaches. Dashed orange line indicates the zero line, where the two measurement approaches have no differences. If the blue line is located above the orange line, then we can conclude that the parents overestimated tummy time recall compared with direct observation, and vice versa. Dashed red lines indicate the upper and lower 95% limits of agreement, computed as the mean difference (bias) plus or minus 1.96 times its SD.

observation. The systematic bias in parent recall for all infants, was 10.40 min/day ( $SD = 25.38$  min), 95% CI [0.74, 20.05], with 95% limits of agreement from  $-39.35$  min to 60.15 min. When the average tummy time duration, as observed directly and reported by parents, was below 15 min/day, the individual differences exhibited lower variance, indicating a higher precision between the approaches. Conversely, at more than 15 min/day, the variance was larger, indicating reduced precision as the average daily tummy time increased.

We conducted an LMM analysis to assess the mean absolute differences between measurement approaches, with birth status as a moderator. The LMM permits the use of all available data (Walker et al., 2019) such that all infants with at least one of the two measures were included in the model ( $n = 31$ ). The final model confirmed a significant Measurement Approach  $\times$  Birth Status interaction,  $F(1, 29.5) = 4.38, p = .04$ , which suggests that tummy time durations obtained using parent recall and direct observation varied between full-term and preterm infants. Additionally, there was a significant main effect for the measurement approach,  $F(1, 29.5) = 8.26, p = .0007$ ; indicating that the tummy time recall differed depending on the type of measurement approach used. The least square means of the tummy time recall stratified by birth status are reported in Table 2.

Pairwise comparisons showed that parents of full-term infants overestimated tummy time recall by an average of 3.56 min/day ( $SE = 5.58$ ), 95% CI = [7.92,  $-14.94$ ], compared with direct observation. However, this difference was not statistically significant ( $p = .53$ ). In contrast, parents of preterm infants overestimated tummy time recall by an average of 22.39 min/day ( $SE = 7.06$ ), 95% CI [ $-7.95, -36.82$ ], compared with direct observation, and this difference was statistically significant ( $p = .003$ ). On the basis of these absolute differences, the percent accuracy of parent recall was 85.79% for full-term infants and 44.63% for preterm infants.

## Discussion

Parent recall is the most frequently used method for assessing tummy time in infants at home. This is the first study to assess the agreement between parent recall for tummy time and the gold-standard direct observation in both full-term and preterm infants within home settings. Consistent with previous

research on parent reports, our findings reported an overall moderate correlation between parent recall and direct observation. Parents of preterm infants tended to significantly overestimate tummy time duration across a 3-day recall period, with the percentage accuracy of the recall being less than 50%.

Among the 29 infants included in our study, 44.4% of full-term infants (8 of 18) and 72.7% of preterm infants (8 of 11) engaged in less than 15 min of tummy time per day. Although parents were requested to record all play sessions on video, the extent of their compliance was voluntary. Thus, infants might have participated in additional tummy time during caregiving activities or at moments when parents were not actively recording. Typically, behavioral studies show an increase in desired behavior because of biases (Althubaiti, 2016). However, our findings suggest that this may not be the case. It is plausible that the observed durations reflect infants' daily routine or even exceed their usual tummy time. Regardless, over 70% of preterm infants in our study were receiving less tummy time than recommended by the AAP. This aligns with a previous population-based study (Zachry & Kitzmann, 2011) and supports Bartlett and Fanning's (2003) findings that preterm infants often prefer play positions other than prone.

Existing research on the correlation between parent recall measures and gold-standard methods of physical activity has focused mainly on toddlers and older children, with limited studies in infants (Arts et al., 2022; Chinapaw et al., 2010). In our study, we found slightly higher correlations between parent recall and direct observation for tummy time compared with those found in physical activity studies of toddlers. This difference is expected because physical activity is a broader concept and more challenging to measure than a specific play position such as tummy time. Additionally, tummy time predominantly occurs indoors under adult supervision (American Academy of Pediatrics, 2016), which allows for better attention to its occurrence. Our results are consistent with those of another study that compared tummy time questionnaires and accelerometers in age-matched infants and reported moderate correlations ( $r = .60$ ; Zhang et al., 2022).

We found that, regardless of the birth status, parents tended to overestimate tummy time on a 3-day recall with an overall a mean systematic bias of 10 min/day and wide 95% limits of agreement, indicating lower precision of parent recall measures. Bland-Altman

**Table 2. Least Square Mean Estimates From the Linear Mixed-Model Analysis for Tummy Time Duration Measured Using Parent Recall and Direct Observation (in Minutes)**

Measurement Approach	Full-Term Infants ( $n = 19$ )			Preterm Infants ( $n = 12$ )		
	Estimate	SE	95% CI	Estimate	SE	95% CI
Direct observation	24.7	5.36	[13.90, 35.49]	15.48	6.48	[2.44, 28.54]
Parent recall	28.21	5.48	[17.18, 39.24]	37.87	6.95	[23.93, 51.83]

Note. CI = confidence interval.

plots showed that the systematic bias tended to increase when tummy time exceeded 15 min/day. Because of limited sample sizes, systematic biases per birth status were not calculated (Giavarina, 2015). However, examining the mean absolute differences we found that parents of preterm infants overestimated tummy time by approximately 22 min/day (44.63% accuracy). When evaluated against our predetermined acceptable accuracy, parent-recall measures for preterm infants may not be acceptable clinically. In contrast, parents of full-term infants overestimated tummy time by 4 min/day (85.79% accuracy), indicating a clinically acceptable agreement. This finding is significant because it suggests that prematurity may affect the statistical and clinical accuracy of parent-recall measures for tummy time in infants.

Because of the limited number of preterm infants in our study, we completed a post hoc descriptive analysis to examine how individual parent recall accuracy varied. We found that 8 of the 11 parents (73%) of preterm infants overestimated tummy time by more than double the directly observed values. Information regarding race, ethnicity, and annual income was unavailable for all 8 parents (3 parents declined to provide this information). As a result, we were unable to conduct a descriptive analysis involving these factors. Nonetheless, it was observed that, among the 8 parents, 5 had circumstances such as being employed full-time, being single parents, or having multiple children. These demographic factors all have the potential to add to the workload of reporting for parents, and their impact should be investigated more thoroughly in future studies using a larger sample size. Finally, building on our findings from the Bland–Altman plots, we conducted a correlation analysis between the duration of directly observed data and the discrepancies between parent recall and direct observation. Our analysis revealed no significant correlation between these values ( $\rho = .22, p = .50$ ). This suggests that the accuracy of parents' recall regarding their preterm infants' tummy time behavior might not be influenced by the volume of data recorded daily but, rather, by the availability of time to observe and accurately document these activities.

Preterm birth also poses significant challenges for families, particularly primary caregivers (Menon, 2012), and prior studies in clinical populations consistently demonstrate the strong influence of caregiver burden on parent-report measures. For example, Corder et al. (2012) found that highly involved parents tend to overestimate their child's physical activity duration. This suggests that parents of preterm infants in our study may have included their own participation or supervision efforts during tummy time, resulting in overestimations. Another potential explanation for the observed overestimation in parents of preterm infants could be attributed to *social desirability*, which is the tendency to exhibit behavior or desire outcomes that are socially favorable, and a common source of error

in proxy reports, particularly among at-risk children (Bornstein et al., 2015). A meta-analysis revealed that 50% of parents of overweight or obese children underreported their children's weight in recall questionnaires despite being aware of the associated health risks (Lundahl et al., 2014). Similarly, in our study, parents of preterm infants may have had a desire to see their child performing well in gross motor activities, which, when coupled with the reporting requirement, could contribute to the overestimation of tummy time durations. In summary, these results underscore the importance of considering birth status when utilizing parent recall as a measure of tummy time in infants.

## Limitations and Future Directions

Our sample size, particularly for preterm infants, is limited compared with those in previous parent report agreement studies (Bennetts et al., 2016). Furthermore, because of this restricted sample size, we did not examine the potential impact of demographic variables on the precision of parental recall. Future studies should build on this initial evidence by recruiting a larger, more diverse infant sample. Considering the potential influence of specific demographic variables on the precision of reporting within our preterm participant group, future studies should account for these variables in their analyses. This would enable a comprehensive assessment of whether premature birth affects the accuracy of tummy time recall.

Another limitation is the potential bias that may arise from parents' awareness of the study focus and their informed knowledge of the parent recall measure, which may have increased recall accuracy. To minimize response bias, our parent recall survey was designed as a "one-attempt" only, discouraging repeated changes in responses.

Despite the smaller sample size, a notable strength of our study is the use of a rigorous gold-standard comparison tool. Through direct observation and frame-by-frame behavioral coding of play positions over 3 days in the home environment, we enhanced the ecological validity of our findings (Franchak, 2019). This study is among the first to validate tummy time measures specifically for preterm infants in the home environment.

## Implications for Occupational Therapy Practice

Tummy time is an important occupational activity for young infants, and accurate assessment of tummy time performance is crucial to improve health outcomes. The findings have the following implications for occupational therapy practice.

- For healthy, full-term infants, a 3-day parent recall measure can be used to monitor tummy time positioning practices and their impact on health indicators cost-effectively.



- For preterm infants, recall measures for tummy time may not be accurate. Objective measures such as wearable sensors should be explored to track tummy time precisely and assist caregivers in promoting tummy time in this group.
- Given the significant differences between recall and direct observation, these measures may not be suitable for clinical documentation of intervention effectiveness. Clinicians and researchers should consider using dense subjective reporting methods (Franchak, 2019) or objective measures (Hewitt et al., 2019) to better capture changes in tummy time after interventions.

## Conclusion

This study presents preliminary evidence to inform the selection of suitable assessment measures for measuring tummy time durations in infants. A 3-day parent recall method shows an acceptable clinical accuracy with direct observation, making it a viable option for tracking tummy time adherence in healthy, full-term infants in clinical practice. The caregiver burden experienced during early months of care for preterm infants may moderate the accuracy of parent reports, suggesting the need for alternative tummy time assessments for this population. To establish the generalizability of these findings, it is imperative to replicate this study with a larger and more diverse sample of preterm infants. 🏠

## Acknowledgments

We extend our gratitude to the participating infants and their families. We thank Kayla Bowler, Jessica Manning, and Dhriti Thakur for their time and dedication in data coding. This research was supported by the Center for Smart Use of Technologies to Assess Real-world Outcomes (C-STAR) pilot project grant (principal investigator: Virginia Chu) and a graduate school dissertation assistantship awarded to Ketaki Inamdar during the final year of her doctoral studies at Virginia Commonwealth University. Selected findings from this study were presented at American Occupational Therapy Association INSPIRE, 2023.

## References

Akoglu, H. (2018). User's guide to correlation coefficients. *Turkish Journal of Emergency Medicine*, 18, 91–93. <https://doi.org/10.1016/j.tjem.2018.08.001>

Althubaiti, A. (2016). Information bias in health research: Definition, pitfalls, and adjustment methods. *Journal of Multidisciplinary Healthcare*, 9, 211. <https://doi.org/10.2147/JMDH.S104807>

American Academy of Pediatrics. (2016). SIDS and other sleep-related infant deaths: Updated 2016 recommendations for a safe infant sleeping environment. *Pediatrics*, 138, e20162938. <https://doi.org/10.1542/peds.2016-2938>

Arts, J., Gubbels, J. S., Verhoeff, A. P., Chinapaw, M. J. M., Lettink, A., & Altenburg, T. M. (2022). A systematic review of proxy-report

questionnaires assessing physical activity, sedentary behavior and/or sleep in young children (aged 0–5 years). *International Journal of Behavioral Nutrition and Physical Activity*, 19, 1–39. <https://doi.org/10.1186/s12966-022-01251-x>

Bartlett, D. J., & Fanning, J. E. K. (2003). Relationships of equipment use and play positions to motor development at eight months corrected age of infants born preterm. *Pediatric Physical Therapy*, 15, 8–15. <https://doi.org/10.1097/01.PEP.0000051693.49199.41>

Bennetts, S. K., Mensah, F. K., Westrupp, E. M., Hackworth, N. J., & Reilly, S. (2016). The agreement between parent-reported and directly measured child language and parenting behaviors. *Frontiers in Psychology*, 7, 1710. <https://doi.org/10.3389/fpsyg.2016.01710>

Bland, J. M., & Altman, D. G. (1999). Measuring agreement in method comparison studies. *Statistical Methods in Medical Research*, 8, 135–160. <https://doi.org/10.1177/096228029900800204>

Bornstein, M. H., Putnick, D. L., Costlow, K. M., & Suwalsky, J. T. D. (2020). Retrospective report revisited: Long-term recall in European American mothers moderated by developmental domain, child age, Person, and metric of agreement. *Applied Developmental Science*, 24, 242–262. <https://doi.org/10.1080/10888691.2018.1462090>

Bornstein, M. H., Putnick, D. L., Lansford, J. E., Pastorelli, C., Skinner, A. T., Sorbring, E., . . . Oburu, P. (2015). Mother and father socially desirable responding in nine countries: Two kinds of agreement and relations to parenting self-reports. *International Journal of Psychology*, 50, 174–185. <https://doi.org/10.1002/ijop.12084>

Carson, V., Lee, E. Y., Hewitt, L., Jennings, C., Hunter, S., Kuzik, N., . . . Tremblay, M. S. (2017). Systematic review of the relationships between physical activity and health indicators in the early years (0–4 years). *BMC Public Health*, 17(Suppl. 5), 854. <https://doi.org/10.1186/s12889-017-4860-0>

Chinapaw, M. J. M., Mokkink, L. B., Van Poppel, M. N. M., Van Mechelen, W., & Terwee, C. B. (2010). Physical activity questionnaires for youth: A systematic review of measurement properties. *Sports Medicine*, 40, 539–563. <https://doi.org/10.2165/11530770-000000000-00000>

Corder, K., Crespo, N. C., van Sluijs, E. M. F., Lopez, N. V., & Elder, J. P. (2012). Parent awareness of young children's physical activity. *Preventive Medicine*, 55, 201–205. <https://doi.org/10.1016/j.ypmed.2012.06.021>

Dudek-Shriber, L., & Zelazny, S. (2007). The effects of prone positioning on the quality and acquisition of developmental milestones in four-month-old infants. *Pediatric Physical Therapy*, 19, 48–55. <https://doi.org/10.1097/01.ppe.0000234963.72945.b1>

Franchak, J. M. (2019). Changing opportunities for learning in everyday life: Infant body position over the first year. *Infancy*, 24, 187–209. <https://doi.org/10.1111/inf.12272>

Giavarina, D. (2015). Understanding Bland Altman analysis. *Biochemia Medica*, 25, 141–151. <https://doi.org/10.11613/BM.2015.015>

Harris, P. A., Taylor, R., Minor, B. L., Elliott, V., Fernandez, M., O'Neal, L., . . . Duda, S. N. (2019). The REDCap consortium: Building an international community of software platform partners. *Journal of Biomedical Informatics*, 95, 103208. <https://doi.org/10.1016/j.jbi.2019.103208>

Harvey, E. A., Fischer, C., Weieneth, J. L., Hurwitz, S. D., & Sayer, A. G. (2013). Predictors of discrepancies between informants' ratings of preschool-aged children's behavior: An examination of ethnicity, child characteristics, and family functioning. *Early Childhood Research Quarterly*, 28, 668–682. <https://doi.org/10.1016/j.ecresq.2013.05.002>

Hesketh, K. D., Crawford, D. A., Abbott, G., Campbell, K. J., & Salmon, J. (2015). Prevalence and stability of active play, restricted movement and television viewing in infants. *Early Child Development and Care*, 185, 883–894. <https://doi.org/10.1080/03004430.2014.963066>

Hewitt, L., Stanley, R. M., Cliff, D., & Okely, A. D. (2019). Objective measurement of tummy time in infants (0–6 months): A validation

- study. *PLoS One*, 14, e0210977. <https://doi.org/10.1371/journal.pone.0210977>
- Kippe, K., Marques, A., Martins, J., & Ligestad, P. A. (2022). Parents' inadequate estimate of their children's objectively physical activity level. *Children*, 9, 392. <https://doi.org/10.3390/children9030392>
- Lee, L. L.-Y. (2015). Fostering upper extremity motor development with an infant prone to play program using an evidence-based approach [Doctoral dissertation, Boston University]. Boston University Libraries. [https://open.bu.edu/bitstream/handle/2144/15662/Lee\\_bu\\_0017E\\_10906\\_web.pdf?sequence=8&isAllowed=y](https://open.bu.edu/bitstream/handle/2144/15662/Lee_bu_0017E_10906_web.pdf?sequence=8&isAllowed=y)
- Lorne, E., Diouf, M., de Wilde, R. B. P., & Fischer, M. (2018). Assessment of interchangeability rate between 2 methods of measurements: An example with a cardiac output comparison study. *Medicine*, 97, e9905. <https://doi.org/10.1097/MD.0000000000009905>
- Lundahl, A., Kidwell, K. M., & Nelson, T. D. (2014). Parental underestimates of child weight: A meta-analysis. *Pediatrics*, 133, e689–e703. <https://doi.org/10.1542/peds.2013-2690>
- Menon, R. (2012). Preterm birth: A global burden on maternal and child health. *Pathogens and Global Health*, 106, 139–140. <https://doi.org/10.1179/20477312X13462106637729>
- Palmer, C. F., Rindler, D., & Leverone, B. (2019). Moving into tummy time, together: Touch and transitions aid parent confidence and infant development. *Infant Mental Health Journal*, 40, 277–288. <https://doi.org/10.1002/imhj.21771>
- Patel, N., & Desai, S. (2020). ABC of face validity for questionnaire. *International Journal of Pharmaceutical Sciences Review and Research*, 65, 164–168. <https://doi.org/10.47583/ijpsrr.2020.v65i01.025>
- Piper, M. C., & Darrah, J. (1994). *Motor assessment of the developing infant*. Saunders.
- Reese, E., & Read, S. (2000). Predictive validity of the New Zealand MacArthur Communicative Development Inventory: Words and Sentences. *Journal of Child Language*, 27, 255–266. <https://doi.org/10.1017/S0305000900004098>
- Ricardo, L. I. C., Da Silva, I. C. M., Martins, R. C., Wendt, A., Gonçalves, H., Hallal, P. R. C., & Wehrmeister, F. C. (2018). Protocol for objective measurement of infants' physical activity using accelerometry. *Medicine and Science in Sports and Exercise*, 50, 1084–1092. <https://doi.org/10.1249/MSS.0000000000001512>
- Russell, D., Kriel, H., Joubert, G., & Goosen, Y. (2009). Prone positioning and motor development in the first 6 weeks of life. *South African Journal of Occupational Therapy*, 39, 11–14.
- Senju, A., Shimono, M., Tsuji, M., Suga, R., Shibata, E., Fujino, Y., . . . Kusunohara, K. (2018). Inability of infants to push up in the prone position and subsequent development. *Pediatrics International*, 60, 811–819. <https://doi.org/10.1111/ped.13632>
- Sommer, I., Nußbaumer-Streit, B., & Gartlehner, G. (2021). WHO-Leitlinie: Körperliche Aktivität, Bewegungsmangel und Schlaf für Kinder unter 5 Jahren [WHO guideline: Physical activity, sedentary behavior and sleep for children under 5 years of age]. *Gesundheitswesen*, 83, 509–511. <https://doi.org/10.1055/a-1489-8049>
- Uzark, K., Smith, C., Yu, S., Lowery, R., Tapley, C., Romano, J. C., & Butcher, J. (2021). Evaluation of a “tummy time” intervention to improve motor skills in infants after cardiac surgery. *Cardiology in the Young*, 32, 1210–1215. <https://doi.org/10.1017/S1047951121003930>
- Van Vlimmeren, L. A., van der Graaf, Y., Boere-Boonekamp, M. M., L'Hoir, M. P., Helders, P. J. M., & Engelbert, R. H. H. (2008). Effect of pediatric physical therapy on deformational plagiocephaly in children with positional preference: A randomized controlled trial. *Archives of Pediatrics and Adolescent Medicine*, 162, 712–718. <https://doi.org/10.1001/archpedi.162.8.712>
- Van Zyl, Z., Maslin, K., Dean, T., Blaauw, R., & Venter, C. (2016). The accuracy of dietary recall of infant feeding and food allergen data. *Journal of Human Nutrition and Dietetics*, 29, 777–785. <https://doi.org/10.1111/jhn.12384>
- Walker, E. A., Redfern, A., & Oleson, J. J. (2019). Linear mixed-model analysis to examine longitudinal trajectories in vocabulary depth and breadth in children who are hard of hearing. *Journal of Speech, Language, and Hearing Research*, 62, 525–542. [https://doi.org/10.1044/2018\\_JSLHR-L-ASTM-18-0250](https://doi.org/10.1044/2018_JSLHR-L-ASTM-18-0250)
- Zachry, A. H., & Kitzmann, K. M. (2011). Caregiver awareness of prone play recommendations. *American Journal of Occupational Therapy*, 65, 101–105. <https://doi.org/10.5014/ajot.2011.09100>
- Zhang, Z., Predy, M., Kuzik, N., Hewitt, L., Hesketh, K. D., Pritchard, L., . . . Carson, V. (2022). Validity of an infant tummy time questionnaire and time-use diary against the GENEActiv Accelerometer. *Measurement in Physical Education and Exercise Science*, 26, 27–38. <https://doi.org/10.1080/1091367X.2021.1941033>

---

**Ketaki Inamdar, PT, PhD**, is Assistant Professor, Department of Human Performance-Physical Therapy, West Virginia University, Morgantown; [ketaki.inamdar@hsc.wvu.edu](mailto:ketaki.inamdar@hsc.wvu.edu)

**Stacey C. Dusing, PhD, PT, FAPTA**, is Sykes Family Chair in Pediatric Physical Therapy, Pediatric Health and Development, and Associate Professor, Division of Biokinesiology and Physical Therapy, University of Southern California, Los Angeles.

**Leroy Thacker, PhD**, is Associate Professor, Department of Biostatistics, Virginia Commonwealth University, Richmond.

**Peter E. Pidcoe, PT, DPT, PhD**, is Professor, Department of Physical Therapy, Virginia Commonwealth University, Richmond.

**Sheryl Finucane, PT, PhD**, is Associate Professor, Department of Physical Therapy, Virginia Commonwealth University, Richmond.

**Virginia Chu, PhD, OTR/L**, is Assistant Professor, Department of Occupational Therapy, Virginia Commonwealth University, Richmond.