SCIENTIFIC ARTICLE

Validity and Reliability of the Thumb Grasp and Pinch Assessment for Children After Reconstruction of Congenital Hypoplastic Thumbs

Ida Neergård Sletten, MD, PhD,* Mona Irene Winge, MD,* Camilla Hellevuo, MD,† Anne Birgit Stavenes, OT,* Inger Helen Bolstad, OT,* Jarkko Jokihaara, MD, PhD†‡

Purpose The Thumb Grasp and Pinch Assessment (T-GAP) is a new instrument for evaluating thumb use in children with congenital hypoplastic thumbs. The assessors video-record the children while they perform nine specific activities and score their grasp types using T-GAP. A high T-GAP score indicates more mature grasp patterns. The developers reported the instrument's validity and reliability for index finger pollicization. This study investigated T-GAP's validity and reliability in children with reconstructed hypoplastic thumbs.

Methods Four hand surgeons and two hand therapists from two hospitals rated video clips of 20 Manske type II and IIIa hands twice in 17 patients who performed the T-GAP at least 1 year after opposition transfer and thumb ligament reconstruction. To investigate the validity, we calculated correlation coefficients for T-GAP scores and clinical outcomes, including thumb ROM, grip and pinch strength, and visual analog assessments of thumb function and appearance. To estimate T-GAP's inter- and intrarater reliability, we calculated intraclass correlation coefficients and their 95% confidence intervals (CIs).

Results Thumb Grasp and Pinch Assessment score showed a strong linear correlation ($r = 0.815-0.944$) and a moderate to strong nonlinear correlation ($\rho = 0.527-0.744$) with visual analog scale assessments of thumb function and appearance, respectively; a moderate nonlinear correlation ($\rho = 0.464$) with grip strength; and a moderate nonlinear correlation ($\rho = 0.541$) with thumb MCP joint range of motion. The intraclass correlation coefficient for the interrater reliability was 0.892 (95% CI, 0.768−0.954) in round 1 and 0.898 (95% CI, 0.754−0.959) in round 2, and for intrarater reliability, the mean was 0.882 (95% CI, 0.785−0.980).

Conclusions Thumb Grasp and Pinch Assessment score had a moderate to strong construct validity and a moderate concurrent validity. Both inter- and intrarater reliability was strong.

Clinical relevance This study supports the T-GAP instrument's validity and reliability for assessing functional outcomes in congenital hypoplastic thumb reconstruction. *(J Hand Surg Am. 2023;■(■):1.e1-e8. Copyright © 2023 by the American Society for Surgery of the Hand. Published by Elsevier Inc. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).)*

Key words Congenital thumb hypoplasia, reliability, Thumb Grasp and Pinch Assessment, validity.

Online Exclusive

CREDIBLE EVALUATION OF TREATMENT outcomes is crucial in studies on children with congenital upper limb anomalies (CULA). During the last decades, researchers have increasingly used patient-reported outcome measures (PROMs) to supplement objective measurements of

*From the *Division of Orthopaedic Surgery, Oslo University Hospital, Oslo, Norway; the †Department of Hand Surgery, Tampere University Hospital, Central Hospital, Tampere, Finland; and the ‡Faculty of Medicine and Health Technology, Tampere University, Tampere, Finland.*

Received for publication August 23, 2022; accepted in revised form December 20, 2022.

No benefits in any form have been received or will be received related directly to this article.

Corresponding author: Ida Neergård Sletten, MD, PhD, Division of Orthopedic Surgery, Oslo University Hospital, Postboks 4950 Nydalen, 0424 Oslo, Norway; e-mail: ida.sletten@icloud.com.

0363-5023/23/■■-0001
https://doi.org/10.1016/j.jhsa.2022.12.017

strength, ROM, and hand dexterity. However, children with hypoplastic thumbs often use alternative grip patterns to compensate for their thumbs' lack of strength and stability. Therefore, hand dexterity tests and PROMs for upper extremity might not accurately reflect poor thumb function. The optimal evaluation of the treatment outcome of hypoplastic thumbs should, thus, include tests for thumb use in daily activities. The current lack of a gold standard thumb function test limits the possibility of comparing results from clinical studies because researchers use different functional tests and PROMs. A good instrument for outcome assessment must be valid and reliable, preferably easy to learn, and convenient to perform.

The Thumb Grasp and Pinch Assessment (T-GAP) is a 5-minute test developed to evaluate pollicization outcomes in children aged 18 months to 18 years.[1] Assessors video record the children's hands performing nine age-dependent activities so that they can either score the test at once or watch the video recording later. The instrument can be used to evaluate pre- and postoperative thumb function in candidates for pollicization or thumb reconstruction and to follow-up with patients over time.[2] The developers use it when they assess children with pollicized thumbs for further interventions, for example, opposition transfer, stabilization procedures, and nonsurgical therapy.[2]

The T-GAP research group has published their findings on the instrument's validity and reliability in children after index pollicization.[1,2] This study aimed to investigate the validity and reliability of the T-GAP after thumb reconstruction in children with congenital thumb hypoplasia.

MATERIALS AND METHODS

We conducted and reported this methodological study following the Quality Appraisal of Diagnostic Reliability tool and Standards for Reporting of Diagnostic Accuracy guidelines.[3,4] The data protection officer of (Oslo University Hospital) approved the research protocol on June 11, 2019. We conducted the study according to the Helsinki declaration. All the patients' caregivers signed a written consent form.

Patients

The study sample was a consecutive subset of children included in an ongoing clinical cohort study of reconstructed congenital hypoplastic thumbs in (Oslo University Hospital). The specific inclusion criteria were as follows: (1) hypoplastic thumb reconstruction with an opposition transfer and if indicated, reconstruction of the collateral ligaments in the thumb MCP joint using the superficial digital flexor from the ring finger,[5,6] (2) minimum 1-year postoperative follow-up time, and (3) patient age of minimum 5 years at follow-up.

We identified the cohort patients by an electronic search in our patient records on the International Classification of Diseases 10th revision diagnosis code Q71.4 and a manual search in our operation log books. We operated on the first patient with the described reconstruction technique in April 2009 and included data of all patients until November 2020 (24 children and 27 hands). All 24 patients were invited to participate in the study; however, three of them declined. Four other patients were not included in this methodological T-GAP part of the study because of their age being under 5 years or follow-up time of less than 1 year. The included 20 hands of 17 patients in this study represented a consecutive series of reconstructed hypoplastic thumbs with video clips of the T-GAP test collected on the same day as the other clinical outcomes (September 2020—October 2021). We evaluated six hands in six girls and 14 hands in 11 boys, of which 10 were in the 5—7-year age group and 10 were in the 8—18-year age group. We classified nine hands as Manske type II and 11 hands as Manske type IIIa before the surgery.[7]

Thumb Grasp and Pinch Assessment

We recorded the children while they performed the T-GAP and used test kit items as instructed by the T-GAP developers,[1,2] with the two sets of activities for the age range of 5—7 years and 8—18 years (Table 1). The raters studied the T-GAP scoring and administration manual, test setup details, scoring sheets, photos illustrating standard and variations in grasp patterns, and two slide show presentations with illustrative video clips. The raters discussed the scoring system and the developers' video clips with the other raters before the start of the study.

The raters evaluated study video recordings in two rounds and scored the children's dominant grasp type for each activity (value, 0—7 points), resulting in a final T-GAP score (0—63). The seven grasp types (and corresponding values) evaluated were as follows: no grasp (0 points), palmar grasp (1 point), ulnar scissor grasp (2 points), radial scissor grasp (3 points), cylindrical grasp (4 points), lateral key pinch (5 points), tip pinch (6 points), and tripod pinch (7 points). Higher T-GAP scores indicated more mature grasp patterns, and scores of 4—7 points noted the use of the thumb. The raters chose the highest score if two grasp patterns were used equally in an activity.

TABLE 1. Thumb Grasp and Pinch Assessment (T-GAP) for Age Groups 5−7 and 8−18 years

Activity	Age 5−7 years	Age 8−18 years
1 Tip pinch	Pick up three pennies one at a time and release into a piggy bank	Thread five plastic beads onto a zip tie
2 Lateral key pinch	Turn a key to open a padlock	
3 Small grasp	Pull cap off a small diameter Crayola marker	Remove cap from ballpoint pen
4 Medium grasp	Turn end of kaleidoscope three times	Hold a paper tube with rubber band and look through the end
5 Large grasp	Twist cap from a peanut butter jar	
6 Manipulation	Form Play to doh into a bowl	Rotate a pencil three times in a handheld pencil sharpener
7 Resistance	Pull back foam pull on slingshot	
8 School	Color inside a circle with a crayon	Write name with a pencil
9 Activity of daily life	Tie shoelaces into a knot	Tie shoelaces into a bow

Activities for the age group from 18 months to 4 years are not shown in the table, as we had no participants in this age group.

They were allowed to watch video recordings as often as they wanted and use pause and slow motion. In addition, they scored how difficult it was to use the T-GAP scale for each patient on a numeric rating scale (NRS difficulty, 0−10). We labeled the minimum value (0 points) as "very easy" and the maximum value (10 points) as "very difficult" with no definition of the intermediate values. We randomized the patient sequence in round 2 (R2) 6 weeks after round 1 (R1). Two raters mistakenly performed R2 without randomization. The raters did not discuss the scorings starting R1 until they had completed R2. They did not review the videos between R1 and R2 or review their scores in R1 while performing R2.

Other subjective outcome measures

We included the parent-proxy version of the Patient-Reported Outcomes Measurement Information System v2.0 Physical Function Upper Extremity Short Form 8a (PROMIS UE 8)[8] and four visual analog scale (VAS) assessments (0−100, 100 = best) of thumb function and appearance: (1) "works like a thumb," (2) "looks like a thumb," (3) "how often do you/does the child use the thumb to pinch versus scissor pinch for small objects," and (4) "how often do you/does the child incorporate the thumb when holding larger objects like a bottle."[9,10] Patients aged >8 years (n = 8), all caregivers (n = 20), one surgeon (I.N.S. or M.I.W.), and one hand therapist (HT, A.B.S.) answered VAS assessments 1 and 2, whereas only the patients and caregivers answered VAS assessments 3 and 4.

Other objective outcome measures

We included the following objective outcomes: patient age at follow-up; Manske type of thumb hypoplasia;[7] active thumb ROM; Kapandji score;[11] grip strength; tip, key (lateral), and tripod (palmar) pinch strength; and the sticker test outcome.[9,10] The HT measured thumb palmar abduction and adduction with the Pollexograph,[12,13] retropulsion as millimeters of thumb lift-off, and radial abduction as the maximal angle between metacarpals 1 and 2, the latter two when the hand was flat on a table. The HT measured flexion and extension in the thumb IP and MCP joints using a goniometer and strength using a calibrated digital Jamar dynamometer and pinch gauge. We reported grip strength as a percentage of normative values in the dominant or nondominant hand.[14] When we reported pinch strength, left-hand values were used as references for nondominant hands because McQuiddy et al[15] gave right/left values instead of dominant and nondominant values.

Raters

Four raters were hand surgeons (I.N.S. or M.I.W.) from two university hospitals with 6−14 years of experience in CULA surgery, and two raters were HTs (A.B.S.) from one university hospital with 12−20 years of experience in evaluating and treating children with CULA.

Statistical methods

We investigated the construct validity of the T-GAP by testing if T-GAP scores and VAS assessments of thumb function and appearance were correlated.[1] We investigated the concurrent validity by testing the correlation between T-GAP scores and objective measures.[1] In addition, we investigated if T-GAP scores correlated to PROMIS UE 8 and NRS difficulty and if NRS difficulty correlated to patient age.

We created scatter plots for the T-GAP score for each patient (the mean of six T-GAP scores in R1) versus patient age, Manske type, PROMs, strength, ROM, and NRS difficulty (the mean of six NRS difficulty assessments of T-GAP scoring both from R1 and R2). We tested associations by two-tailed bivariate correlations. Mean T-GAP scores for each patient followed the normal distribution in R1 and R2. For normally distributed clinical outcome scale data, we calculated both Pearson's correlation coefficient (r) for linear correlations and Spearman's rank correlation coefficient (ρ) for nonlinear correlations. We calculated only ρ for the nonnormal clinical outcome scale and ordinal data. Interpretation of correlation coefficients depends on the study setting. We used the same criteria as the T-GAP developers: r and ρ values below 0.4 as low correlation, 0.41−0.69 as moderate correlation, and above 0.70 as strong correlation.[1]

To estimate reliability for the nine activities and the T-GAP score, we calculated average intraclass correlation coefficient (ICC) estimates and their 95% confidence intervals (CIs) based on an absolute agreement, two-way mixed model.[16] Interpretation of ICC estimates is also dependent on the study setting. The guidelines commonly applied for kappa statistics are also often applied to ICCs.[17] To compare our results, we used the same, slightly stricter cut-off values as the T-GAP developers: we regarded ICC estimates over 0.9 as excellent, 0.75−0.89 as strong, 0.6−0.74 as moderate, 0.4−0.59 as fair, and below 0.40 as poor.[2] ICC estimates that were not significantly different from 0 implied random agreement. We set the significance level to 0.05. The only missing data were one patient's two VAS assessments by the HT.

RESULTS

We examined the 20 hands with a median of 4 years (range, 1−11) after surgery. The average of the means of T-GAP scores was 44.7 (95% CI, 42.5−46.9) in R1 and 43.9 (95% CI, 41.6−46.2) in R2. The total ranges in T-GAP scores for each patient among the six raters followed the normal distribution, and the mean was 8.5 (95% CI, 6.9−10.2) in R1 and 9.5 (95% CI, 8.3−10.6) in R2.

Validity

The T-GAP scores in R1 correlated positively to VAS assessments of the thumb function and appearance, ROM in the thumb MCP joint, and grip strength (Table 2). T-GAP scores did not correlate to patient age, Manske classification, or PROMIS UE8 (Table 2).

The NRS difficulty of scoring T-GAP was a mean of 4.0 (95% CI, 3.5−4.4) in R1 and a mean of 3.8 (95% CI, 3.4−4.2) in R2. NRS difficulty score for each patient did not correlate to their T-GAP score in either R1 or R2 but correlated in R1 to the raters' range in T-GAP scores for each patient ($r = 0.767$, $P < 0.05$; $\rho = 0.798$, $P < 0.05$) and to patient age ($\rho = -0.682$, $P < 0.05$).

Reliability

The interrater reliability for the total T-GAP score and individual scoring for all but one of the nine activities was strong to excellent in both rounds (Table 3). For scoring activity 3 (small grasp), the ICC estimates implied moderate interrater reliability, with wide CIs ranging from poor to excellent. In activity 3, one of the HT raters scored 1 point (palmar grasp only) for more than half of the hands in both rounds, a score the other five raters rarely used. Interrater reliability for activity 3 (small grasp) for these five raters only was still slightly lower than that for the other activities, with an ICC estimate of 0.695 (95% CI, 0.418−0.864; $P < 0.001$) that indicated moderate (fair to strong) agreement in R1, and an ICC estimate of 0.812 (95% CI, 0.643−0.916; $P < 0.001$) that indicated strong (moderate to excellent) agreement in R2.

The intrarater reliability for the T-GAP score was mean 0.882 (95% CI, 0.785−0.980), which is strong or excellent for all but one of the surgeon raters (Table 4). We have reported intrarater ICC estimates for all nine activities in Supplementary Table 1, available online on the Journal's website at www.jhandsurg.org. The intrarater reliability was also the lowest for activity 3 (small grasp) for four raters. The raters with the longest clinical experience did not have better intrarater reliability for scoring the T-GAP than the less experienced.

DISCUSSION

The correlations between T-GAP scores and VAS assessments of thumb function and appearance, and thumb ROM and strength supported the instrument's validity in evaluating reconstruction outcomes in hypoplastic thumbs. In addition, our findings of strong to excellent reliability for the T-GAP was consistent with the developers' findings in pollicized hands.[2]

TABLE 2. Correlations Between T-GAP Score and Clinical Outcomes

	Linear correlation*			Nonlinear correlation		
Clinical outcome	*r*	*P*	Interpretation	*ρ*	*P*	Interpretation
Patient age at follow-up				0.235	0.32	ns
Manske's classification				−0.380	0.10	ns
VAS "works like a thumb"						
Patient	0.944	<0.05	Strong	0.627	0.10	ns
Caregiver	0.382	0.10	ns	0.225	0.34	ns
Surgeon				0.182	0.44	ns
HT				0.727	<0.05	Strong
VAS "looks like a thumb"						
Patient				0.218	0.60	ns
Caregiver				0.524	<0.05	Moderate
Surgeon				0.108	0.65	ns
HT				0.497	<0.05	Moderate
VAS "how often do you/does the child use the thumb to pinch versus scissor pinch for small objects"						
Patient				0.691	0.06	ns
Caregiver				0.744	<0.05	Strong
VAS "how often do you/does the child incorporate the thumb when holding larger objects like a bottle"						
Patient	0.815	<0.05	strong	0.303	0.47	ns
Caregiver				0.527	<0.05	Moderate
Parent-proxy PROMIS UEF8				0.072	0.76	ns
Thumb IP joint total ROM				0.015	0.95	ns
Thumb MCP joint total ROM				0.541	<0.05	Moderate
Radial abduction	0.274	0.24	ns	0.207	0.38	ns
Palmar abduction	0.178	0.45	ns	0.171	0.47	ns
Palmar adduction				−0.022	0.93	ns
Retropulsion				0.192	0.42	ns
Kapandji score				0.151	0.52	ns
Grip strength (%)	0.421	0.07	ns	0.464	<0.05	Moderate
Pinch (%)						
Tip	0.382	0.10	ns	0.423	0.06	ns
Key	−0.017	0.94	ns	−0.060	0.80	ns
Tripod	0.220	0.35	ns	0.272	0.25	ns
Sticker test				0.047	0.84	ns

HT, hand therapist; NRS, numeric rating scale; ns, correlation coefficient not significantly different from zero at the 0.05 level; PROMIS UEF8, Patient-Reported Outcomes Measurement Information System Physical Function Upper Extremity Short Form 8; *ρ*, Spearman's rank correlation coefficient; *r*, Pearson's correlation coefficient; T-GAP, Thumb Grasp and Pinch Assessment; VAS, visual analog scale.
*Only calculated for normally distributed scale data.

Our study's internal validity was high because the raters scored the hands independently, did not discuss the instrument or the video clips after the study had started, and could not access their scores from R1 during R2. We did not consider the nonrandomization of the patient sequence for two raters in R2 to have caused any impact on the intra-rater reliability. It is unlikely that the raters could have remembered the order of nine individual scores of 20 patients after 6 weeks.

TABLE 3. Interrater reliability for T-GAP score and nine activities

Activity	Round	ICC (95% CI)	Reliability
T-GAP score	1	0.892 (0.768−0.954)	Strong (strong to excellent)
	2	0.898 (0.754−0.959)	Strong (strong to excellent)
1 Tip pinch	1	0.899 (0.800−0.956)	Strong (strong to excellent)
	2	0.935 (0.876−0.971)	Excellent (strong to excellent)
2 Lateral key pinch	1	0.812 (0.651−0.915)	Strong (moderate to excellent)
	2	0.891 (0.797−0.951)	Strong (strong to excellent)
3 Small grasp	1	0.616 (0.319−0.819)	Moderate (poor to strong)
	2	0.621 (0.312−0.823)	Moderate (poor to strong)
4 Medium grasp	1	0.838 (0.699−0.927)	Strong (moderate to excellent)
	2	0.913 (0.838−0.961)	Excellent (strong to excellent)
5 Large grasp	1	0.995 (0.990−0.998)*	Excellent
	2	0.993 (0.987−0.997)†	Excellent
6 Manipulation	1	0.949 (0.903−0.977)	Excellent
	2	0.914 (0.837−0.961)	Excellent (strong to excellent)
7 Resistance	1	0.916 (0.842−0.962)	Excellent (strong to excellent)
	2	0.857 (0.723−0.937)	Strong (moderate to excellent)
8 School	1	0.862 (0.734−0.938)	Strong (moderate to excellent)
	2	0.876 (0.738−0.947)	Strong (moderate to excellent)
9 Activity of daily life	1	0.852 (0.700−0.935)	Strong (moderate to excellent)
	2	0.849 (0.705−0.933)	Strong (moderate to excellent)

CI, confidence interval; ICC, intraclass correlation coefficient; T-GAP, Thumb Grasp and Pinch Assessment.
All ICCs were significant at the $P < 0.05$ level.
*S4 excluded as no variance in the scores for all 20 patients in at least one of the sessions.
†S3 and S4 excluded as no variance in the scores for all 20 patients in at least one of the sessions.

TABLE 4. Intrarater reliability for T-GAP score

Rater	ICC (95% CI)	Reliability
S1	0.932 (0.827−0.973)	Excellent (strong to excellent)
S2	0.698 (−0.166 to 0.922)	Moderate (poor to excellent)
S3	0.901 (0.747−0.961)	Excellent (moderate to excellent)
S4	0.943 (0.860−0.977)	Excellent (strong to excellent)
HT1	0.934 (0.789−0.976)	Excellent (strong to excellent)
HT2	0.886 (0.709−0.955)	Strong (moderate to excellent)

CI, confidence interval; HT, hand therapist; ICC, intraclass correlation coefficient; S, surgeon; T-GAP, Thumb Grasp and Pinch Assessment.
All ICCs were significant at the $P < 0.05$ level.

The external validity was also high. We included a consecutive series of patients with minimal attrition bias and recruited surgeons and HTs as raters from two countries. None of the raters had any previous experience with scoring the T-GAP, and we considered our reliability outcomes applicable to similarly experienced HTs and surgeons. We had neither experience with the T-GAP video recordings nor patient instructions and noted that both were not always optimal. The hands were not clearly visible at all times because some children did not sit still and moved their heads in front of the camera. If more than one assessor gave instructions or the assessor commented on their performance during the test, the children could become confused, distracted, or change their grasp style. We believe the challenges of our first experience also increased the generalizability of the study findings.

The main limitation of our study was the relatively small sample of patients. More patients would have increased the power of the correlation analyses between the T-GAP scores and the clinical outcomes. We had an even distribution of patients in the two oldest of the three age groups. Thus, our results applied to patients between 5 and 18 years.

In 21 pollicized hands, the T-GAP developers reported a moderate correlation between T-GAP scores and ROM and four hand dexterity tests and a strong correlation between T-GAP scores and grip and pinch strength values.[1] In our study, VAS assessments of thumb appearance and function correlated more with T-GAP scores than individual objective measurements of ROM and strength. This finding emphasized the importance of PROMs and functional tests with daily activities in clinical studies because each objective measurement of ROM and strength gives limited information separately.

Unsurprisingly, T-GAP scores did not correlate with PROMIS UE 8, as this is a generic PROM reflecting the whole arm function of both upper extremities with a possible floor effect for children with CULA limited to the hand plate only. In addition, the lack of correlation might have reflected that the grasp pattern choice was irrelevant to the overall function if less mature grasp types worked well for the patients.

The negative correlation between mean NRS difficulty for each patient and patient age in R1 only might have implied a learning curve in scoring the youngest patients. The lack of correlation between mean NRS difficulty and their T-GAP score in both rounds might have reflected the quality of the instrument, because it was neither easier nor more difficult to score the best functioning thumbs. Both the mean NRS difficulty and the T-GAP interrater reliability were slightly better in R2, which might also have implied a general learning curve. We rated the overall difficulty of scoring each patient. It might have been more valuable if we had rated the problem of scoring each activity in each patient instead.

The T-GAP developers reported strong to excellent inter- and intrarater reliability among four raters from two scoring sessions with a 2-week interval in a cohort of 11 pollicized hands.[2] Despite no previous experience with the instrument, we found similarly good reliability. The reliability analyses for each activity within T-GAP gave important information about the instrument. The lowest reliability found in activity 3 suggested that it might benefit from modification. We regarded this as an example of how a good general agreement can be brought down by an outlier rater in a reliability study with few raters. Appropriate education on how to better score activity 3 may overcome our finding of this activity's lower reliability.

Our study supported the value of the T-GAP as an outcome measure in reconstructed hypoplastic thumbs. It gives the assessor an immediate impression of how much a child uses the thumb in everyday activities. That is not always easily detectable unless the assessor observes the child playing with multiple objects. The T-GAP also provides the assessors with a descriptive tool for different grasp types seen in children with CULA, which can standardize clinical practice.

We recommend performing video recordings while standing behind the child, ensuring the best view of the hand. Only one assessor should give verbal instructions throughout the test to maintain the child's focus. The assessor should avoid questions about hand function during the test because the child may change grasp style accordingly. Video clip reruns were particularly beneficial in our study because many children performed the activities quickly and used more than one grasp type per activity.

Owing to the T-GAP's design, normal hands will not obtain a maximum total score as it is natural to use less mature grasps for some activities. We propose collection of normative, age-dependent T-GAP scores for comparison with scores of children with hypoplastic or absent thumbs. We also suggest future comparative studies with pre-and postoperative T-GAP scores for both reconstructed hypoplastic thumbs and pollicized indices.

ACKNOWLEDGMENTS

We wish to thank Wendy Tomhave for all her help in providing, at our request, the necessary Thumb Grasp and Pinch Assessment information for our study. The Thumb Grasp and Pinch Assessment administration and scoring forms are now available at www.wendytomhave.com.

REFERENCES

1. Kollitz KM, Tomhave WA, Van Heest AE, Moran SL. A new, direct measure of thumb use in children after index pollicization for congenital thumb hypoplasia. *J Hand Surg Am*. 2018;43(11):978−986.e1.
2. Tomhave WA, Kollitz KM, Moran SL. Inter- and intrarater reliability of the thumb grasp and pinch assessment for children following index pollicization for congenital thumb hypoplasia. *J Hand Surg Am*. 2019;44(7):618.e1−618.e8.
3. Lucas NP, Macaskill P, Irwig L, Bogduk N. The development of a quality appraisal tool for studies of diagnostic reliability (QAREL). *J Clin Epidemiol*. 2010;63(8):854−861.
4. Bossuyt PM, Reitsma JB, Bruns DE, et al. STARD 2015: an updated list of essential items for reporting diagnostic accuracy studies. *BMJ*. 2015;351:h5527.
5. Kozin SH, Ezaki M. Flexor digitorum superficialis opponensplasty with ulnar collateral ligament reconstruction for thumb deficiency. *Tech Hand Up Extrem Surg*. 2010;14(1):46−50.
6. Manske PR, McCarroll HR Jr. Reconstruction of the congenitally deficient thumb. *Hand Clin*. 1992;8(1):177−196.
7. Manske PR, McCarroll HR Jr, James M. Type III-A hypoplastic thumb. *J Hand Surg Am*. 1995;20(2):246−253.

8. Cella D, Yount S, Rothrock N, et al. The Patient-Reported Outcomes Measurement Information System (PROMIS): progress of an NIH roadmap cooperative group during its first two years. *Med Care*. 2007;45(5 Suppl 1):S3−S11.
9. Zlotolow DA, Tosti R, Ashworth S, Kozin SH, Abzug JM. Developing a pollicization outcomes measure. *J Hand Surg Am*. 2014;39(9):1784−1791.
10. Sletten IN, Røkkum M, Winge MI. Outcome after pollicization for congenital thumb deficiency: a cohort study of cases in a single unit, 1987 to 2016. *J Hand Surg Am*. 2022;47(5):479.e1−479.e9.
11. Kapandji A. Clinical test of apposition and counter-apposition of the thumb. *Ann Chir Main*. 1986;5(1):67−73.
12. de Kraker M, Selles RW, Molenaar TM, Schreuders TA, Hovius SE, Stam HJ. Palmar abduction measurements: reliability and introduction of normative data in healthy children. *J Hand Surg Am*. 2009;34(9):1704−1708.
13. de Kraker M, Selles RW, Schreuders TA, Hovius SE, Stam HJ. The Pollexograph: a new device for palmar abduction measurements of the thumb. *J Hand Ther*. 2009;22(3):271−276.
14. Bohannon RW, Wang YC, Bubela D, Gershon RC. Handgrip strength: a population-based study of norms and age trajectories for 3- to 17-year-olds. *Pediatr Phys Ther*. 2017;29(2):118−123.
15. McQuiddy VA, Scheerer CR, Lavalley R, McGrath T, Lin L. Normative values for grip and pinch strength for 6- to 19-year-olds. *Arch Phys Med Rehabil*. 2015;96(9):1627−1633.
16. Hallgren KA. Computing inter-rater reliability for observational data: an overview and tutorial. *Tutor Quant Methods Psychol*. 2012;8(1):23−34.
17. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. 1977;33(1):159−174.