Comparison of Functional Outcome Scores in Radial Polydactyly

Robert R. Dijkman, MD, Christianne A. van Nieuwenhoven, MD, PhD, Ruud W. Selles, PhD, and Steven E.R. Hovius, MD, PhD

Investigation performed at the Department of Plastic and Reconstructive Surgery and Hand Surgery, Rotterdam, the Netherlands

Background: A wide range of outcome assessment systems have been used to describe the results and evaluate residual impairment after surgery for radial polydactyly. We conducted a study to determine which of these assessment systems should be considered superior for the most common types of radial polydactyly (types II and IV).

Methods: Ten outcome assessment systems were selected. Three examiners independently evaluated thirty-seven patients, aged four to twenty-two years, with radial polydactyly. Patients completed two manual activity questionnaires. Interobserver reliability was determined with use of an intraclass correlation coefficient (ICC). Validity was assessed by correlating the results derived with the outcome assessment systems with functional visual analog scale (VAS), aesthetic VAS, and manual activity questionnaire scores.

Results: Thirty-seven patients (forty-one hands with radial polydactyly) were evaluated. All patients were assessed by at least two examiners. Reliability was highest for the Japanese Society for Surgery of the Hand (JSSH), Cheng et al., and Tada et al. assessment systems (overall ICCs $\geq$ 0.70). The JSSH system had the highest overall correlations (r_s ranging from 0.48 to 0.80 and 0.45 to 0.63) with functional and aesthetic VAS scores. No significant correlations were found between the outcome scores and the results of the manual activity questionnaires after an average follow-up time of 112 months.

Conclusions: Interobserver reliability was highest for the JSSH classification, which also showed superior correlations with both examiner-rated and patient-rated VAS scores for functional and aesthetic outcome compared with the other nine assessment systems. The finding of a poor correlation between the outcome scores and the results of manual activity questionnaires is in agreement with findings in published literature. We recommend the JSSH assessment method for the scientific evaluation of the outcomes, in terms of body structure and function, of the treatment of radial polydactyly.

Level of Evidence: Therapeutic Level IV. See Instructions for Authors for a complete description of levels of evidence.

Peer Review: This article was reviewed by the Editor-in-Chief and one Deputy Editor, and it underwent blinded review by two or more outside experts. It was also reviewed by an expert in methodology and statistics. The Deputy Editor reviewed each revision of the article, and it underwent a final review by the Editor-in-Chief prior to publication. Final corrections and clarifications occurred during one or more exchanges between the author(s) and copyeditors.

Radial polydactyly is a relatively common and diverse congenital hand difference[1]. Many reports have described the results and complications of the complex surgical treatment of radial polydactyly. The postoperative evaluation methods applied in these studies have ranged from mainly descriptive systems[2] to strict scoring forms[3].

Disclosure: One or more of the authors received payments or services, either directly or indirectly (i.e., via his or her institution), from a third party in support of an aspect of this work. None of the authors, or their institution(s), have had any financial relationship, in the thirty-six months prior to submission of this work, with any entity in the biomedical arena that could be perceived to influence or have the potential to influence what is written in this work. Also, no author has had any other relationships, or has engaged in any other activities, that could be perceived to influence or have the potential to influence what is written in this work. The complete **Disclosures of Potential Conflicts of Interest** submitted by authors are always provided with the online version of the article.

A commentary by Marybeth Ezaki, MD, is linked to the online version of this article at jbjs.org.

TABLE I Intraclass Correlation Coefficients (ICCs) for the Various Outcome Assessment Systems, for Different Combinations of Examiners ($p < 0.05$)

Examiners	Tada	Ogino	Larsen	Horii	ALURRA	Cheng	JSSH	Wood	Dobyns	Tuch
R.R.D., C.A.v.N., & S.E.R.H.	0.68	0.59	0.66	0.67	0.68	0.70	0.71	0.60	0.56	0.54
R.R.D. & C.A.v.N.	0.62	0.57	0.71	0.64	0.65	0.70	0.77	0.34	0.56	0.45
R.R.D. & S.E.R.H.	0.76	0.47	0.69	0.67	0.58	0.64	0.73	0.57	0.58	0.49
C.A.v.N. & S.E.R.H.	0.82	0.89	0.73	0.76	0.82	0.85	0.75	0.91	0.75	0.70

Early outcome studies of radial polydactyly often provided case-specific results of surgery[4]. In later years, postoperative assessment was structured on the basis of certain evaluation criteria[5,6]. A comprehensive scoring system was first introduced by Tada et al. in 1983[7]. This commonly used[8-12] and frequently modified[13-16] system consists of scores for range of motion, joint instability, and malalignment parameters. Consequently, the definition of outcome is based solely on functional parameters in the Tada scoring system. Other assessment systems include aesthetic and subjective factors as important features of postoperative evaluation[17-19].

Despite the large number of assessment systems and their frequent use, there is no evidence that these systems are reliable and valid indicators of overall outcome or residual impairment after surgical treatment of radial polydactyly or that they can be applied regardless of observer experience. In addition, patient and caregiver satisfaction and manual activity have become increasingly important determinants of outcome in contemporary research, but the relationship between patient satisfaction and manual activity and functional and aesthetic outcome after the treatment of radial polydactyly has not been well described.

The purpose of this study was to determine which outcome assessment system is most suitable for evaluation after the surgical treatment of most common types of radial polydactyly. We evaluated the interobserver reliability of ten outcome assessment systems that assess body structure and function. To assess the validity of these ten outcome assessment systems, we correlated the outcome scores with visual analog scale (VAS) scores for functional and aesthetic results assigned by three examiners and by the patients (or parents). Furthermore, we measured manual activity using questionnaires and correlated their results with those derived with the outcome assessment systems. Although this report contains data describing overall outcome, our methodology was not designed to describe our long-term results of treatment.

Materials and Methods

Patients and Examiners

We examined the patient registry for patients with radial polydactyly treated at our hospital from 1993 to 2011. Patients operated on for Flatt[20] (Wassel[4]) type-II or IV radial polydactyly who were four years of age or older at the time of the follow-up and who had been followed for a minimum of one year postoperatively were eligible for inclusion. Patients with severe psychomotor impairment due to multiple congenital anomalies or perinatal asphyxia were excluded from the study.

Two surgeons experienced in the treatment of congenital hand differences (C.A.v.N. and S.E.R.H.) and one well-trained medical intern (R.R.D.) separately evaluated the patients during their visit to the outpatient clinic. The intern was prepared for this project during several international clerkships with renowned surgeons who treated congenital hand differences and was trained in the performance of measurements for children by a pediatric hand therapist. He was not involved in the surgical treatment of the patients and was considered an independent examiner. This study was done with the approval of an accredited Medical Research Ethics Committee (MEC-2010-295) and in accordance with the Declaration of Helsinki.

Outcome Assessment Systems

Ten clinical assessment systems were selected on the basis of the available literature[5-7,13-15,17-19,21]. In this cross-sectional study, the three examiners were completely blinded with regard to each other's evaluations. Every examiner performed all standardized measurements necessary to calculate the overall score of all outcome assessment systems for all patients, with the exception of the strength measurements, which were conducted only once. If required by the specific assessment system, measurements of the affected hand were compared with those of the unaffected, contralateral hand. With bilateral involvement, normative values were used for comparison.

The content validity of each outcome assessment system was evaluated by consensus. All examiners were separately asked to indicate the seven items essential to assess outcome in patients with radial polydactyly. After the first round, the authors agreed on six and disagreed on two items. In the second round, after they had read each other's comments, the following items were considered, by consensus, to be determinants of content validity of an outcome assessment system for radial polydactyly: thumb alignment, interphalangeal and metacarpophalangeal joint stability, combined interphalangeal and metacarpophalangeal range of motion, appearance of the nail and nail folds, prominence or residual deformity of the pulp, and patient satisfaction (Fig. 1).

Clinical Patient Evaluation

Interphalangeal and metacarpophalangeal joint motions, joint stability, and joint alignment were all measured with use of a handheld goniometer. The Pollexograph[22] and the Kapandji score[23] were used to measure palmar abduction and opposition. Tip, tripod, and key pinch strength were measured with use of a pinch strength dynamometer (Baseline; FEI [Fabrication Enterprises Inc.], White Plains, New York). Thumb length was measured from the palpable base of the proximal phalanx to the tip of the thumb. Girth of the thumb was measured at the interphalangeal joint with the joint in neutral position. Pulp circumference was defined as the circumference at the level of the lunula. The width of the nail was measured at the lateral edges of the nail, where the lateral nail folds meet the hyponychium.

Examiners were asked to estimate whether the first web space and the thumb size were sufficient to allow acceptable function. In addition, examiners and patients (or their caregivers) were asked to assess functional and aesthetic outcome on a VAS[2,8,24,25]. The functional VAS ranged from "no functional thumb at all" to "a perfectly functioning thumb." The aesthetic VAS ranged from "so disfigured that the patient would prefer not to have a thumb at all" to "a perfectly normal looking thumb."

The caregiver VAS score was used only if the child appeared to be too young to understand the question. Most children could independently rate their thumb by the age of eight years. For simplicity, we will refer to these results

	Tada[1]	Ogino[1]	Larsen[1]	Horii[2]	ALURRA[2]	Cheng[2]	JSSH[2]	Wood[3]	Dobyns[3]	Tuch[3]
FUNCTION										
Flex. / Ext.*	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Stability*	✓	✓	✓	✓	✓	✓	✓		✓	✓
Alignment*	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Abduction							✓		✓	
Opposition						✓			✓	✓
Ext. lag							✓	✓		
Strength			✓			✓			✓	✓
1st web						✓				✓
Overall function								✓		
AESTHETICS										
Nail*					✓	✓	✓		✓	
Pulp*					✓	✓	✓		✓	✓
Size					✓		✓		✓	✓
Scar							✓	✓		
Prominence*						✓	✓			✓
Overall appearance				✓				✓		
SUBJECTIVE EVALUATION										
Pain							✓			
Patient satisfaction*							✓			✓
SCORING										
Excellent							✓	✓	✓	n.a.
Good	✓	✓	✓	✓	✓	✓	✓	✓	✓	n.a.
Fair	✓	✓	✓	✓	✓	✓	✓	✓		n.a.
Poor	✓	✓	✓	✓	✓	✓	✓	✓	✓	n.a.
Scale	0-5	-1-5	0-9	0-7	0-24	0-30	0-20	n.a.	n.a.	n.a.
VALIDITY										
Content†	3/7	3/7	3/7	3/7	5/7	6/7	7/7	2/7	5/7	6/7

Fig. 1

Summary of the reviewed outcome assessment systems. 1 = purely functional outcome scores; 2 = outcome scores evaluating function, aesthetics, and subjective aspects; 3 = structured outcome criteria systems; * = essential item based on consensus; † = content validity based on consensus and the number of essential items given; and n.a. = not available.

as "patient-rated" throughout the manuscript. Patients were asked to assess pain on a numeric rating scale with 0 representing no pain at all and 10 representing the worst pain imaginable[26].

Questionnaires

Prior to their visit to the clinic, all patients were asked to fill out the ABILHAND-Kids questionnaire[27]. Additionally, a modified Prosthetic Upper Extremity Functional Index (PUFI) questionnaire[28] was completed by patients with a unilateral congenital hand difference. Both questionnaires quantify the child's manual ability by assessing ease of performance of everyday tasks.

Parents or caregivers of younger patients were asked to verbally or manually assist their child in completing the questionnaires when needed. No hard cutoff age was used to determine the need for parent or caregiver assistance, although most children could independently fill out the questionnaires by the age of ten years.

Source of Funding

This project was funded by a research grant from the NutsOhra foundation. The authors declare an absence of financial conflicts of interest.

Statistical Methods

All outcome scores were calculated after clinical evaluation of the patient. Interobserver reliability of the outcome assessment systems was calculated with an intraclass correlation coefficient (ICC 2, 1) with use of a two-way random single measure agreement model[29]. A Spearman correlation coefficient (r_s) was used to assess the relationship between the scores on the outcome assessment systems and the functional and aesthetic VAS scores. Additionally, a Spearman correlation coefficient was used to analyze correlation between the outcome assessment system scores and the results of the manual activity questionnaires.

Results

We identified 211 patients with radial polydactyly from the patient registry; fifty-eight (27%) of them met the inclusion criteria. Thirty-seven (64% response rate) with a total of forty-one cases of radial polydactyly were seen and examined in the outpatient clinic. Examiner R.R.D. assessed all forty-one cases. C.A.v.N. and S.E.R.H. assessed forty (98%) and twenty (49%) cases, respectively. The average age at the time of follow-up was ten years (range, four to twenty-two years); the average age at the first surgical procedure was 1.9 years (range, less than one to seven years) (see Appendix).

Figure 1 shows the items evaluated in each outcome assessment system. Seven systems (Tada et al.[7], Ogino et al.[13], Larsen and Nicolai1[5], Horii et al.[14], ALURRA [alignment, ulnar and radial stability, range of motion and aesthetical aspects][18], Cheng et al.[17],

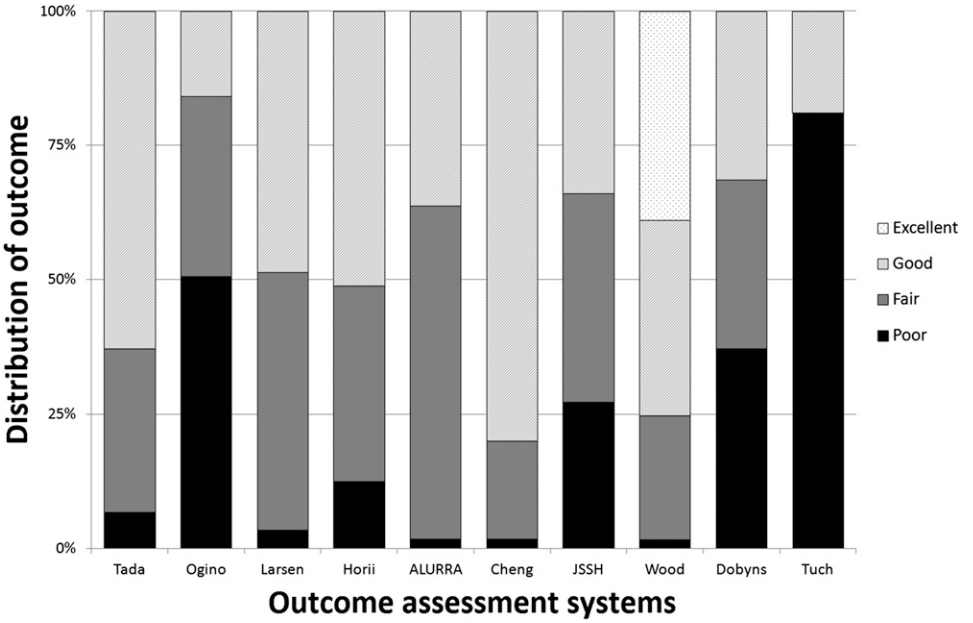

Fig. 2
Distribution of outcomes per system.

and JSSH [Japanese Society for Surgery of the Hand])[19] were designed to calculate an outcome score. The three remaining systems (Wood[21], Dobyns et al.[5], and Tuch et al.[6]) use combined evaluation criteria that define overall outcome. Each system contains three to twelve items. Three systems (Tada, Ogino, and Larsen) do not evaluate aesthetic aspects of outcome. Patient satisfaction is integrated as a determinant of outcome in the JSSH and Tuch systems. Regarding content validity, the JSSH, Cheng, and Tuch assessment systems encompassed the most items defined as essential by the examiners (Fig. 1).

The results of the assessments of all patients are shown for each assessment system in Figure 2. While median outcome scores were fair to good, Figure 2 clearly illustrates that the conclusion regarding overall outcome heavily depends on the

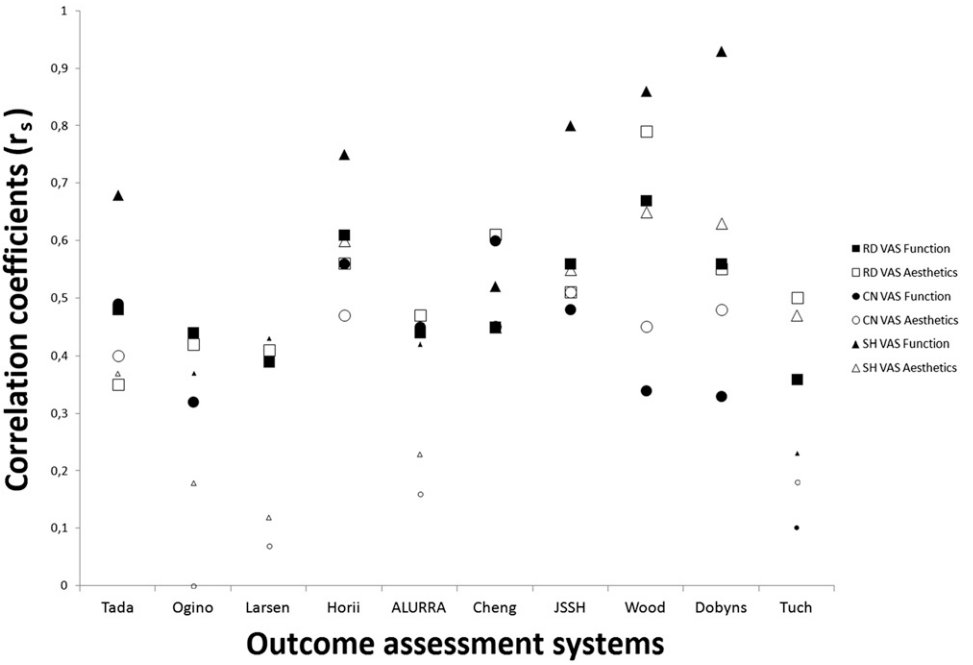

Fig. 3
Correlations (r_s) between each examiner's VAS scores rating functional and aesthetic outcome and the scores derived with each assessment system. The larger markings represent significant correlations ($p \leq 0.05$), and the smaller markings represent nonsignificant correlations ($p > 0.05$).

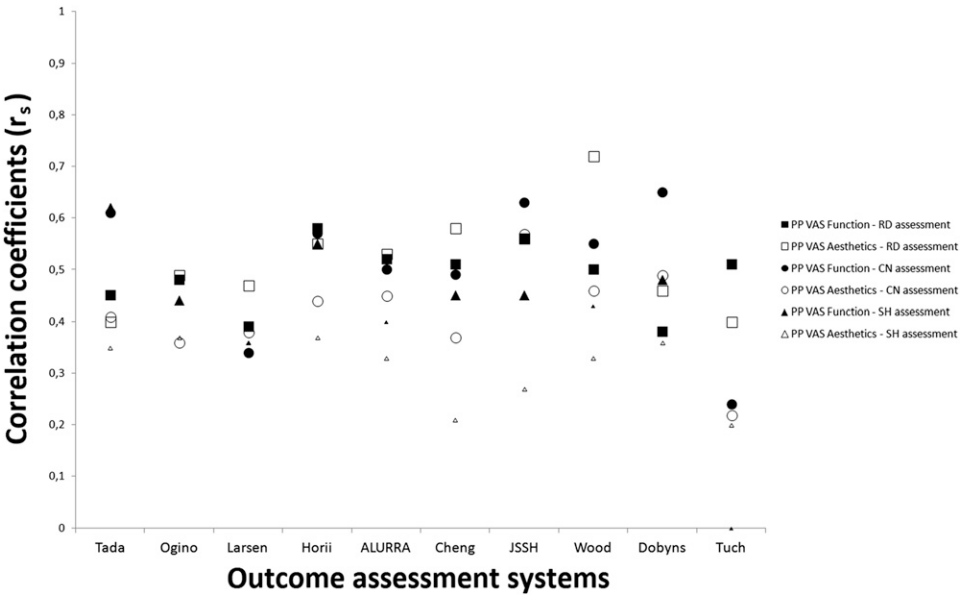

Fig. 4

Correlations (r_s) between patient-rated VAS scores of functional and aesthetic outcome and the scores derived with each assessment system by the different examiners. The larger markings represent significant correlations ($p \leq 0.05$), and the smaller markings represent nonsignificant correlations ($p > 0.05$). PP = patient/parent.

assessment system applied. For example, while approximately 75% of all outcomes were in the "good" to "excellent" range when evaluated with the Cheng and Wood systems, <25% were considered "acceptable" according to the Dobyns system. Figure 2 also presents information on the discriminative properties of the outcome assessment systems, as some systems classify the majority of patients in only one or two categories. Although not shown in Figure 2, a number of the assessment systems were not sensitive enough to enable one or more of the examiners to detect a difference between "poor" and "fair" outcomes. This was the case for the Larsen (examiner C.A.v.N.), ALURRA (examiner S.E.R.H.), Cheng (examiners R.R.D. and C.A.v.N.), and Wood (examiners C.A.v.N. and S.E.R.H.) systems. The Tuch system lacks a "fair" outcome category.

The interobserver reliability of each outcome assessment system is shown in Table I. The overall agreement between the examiners was highest for the JSSH assessment system, followed by the Cheng and Tada systems (overall ICCs ≥ 0.70). The Cheng and Tada systems had, respectively, one and two ICCs of <0.70. Agreement between the hand surgeons was generally higher than that between either of them and the independent examiner.

Patients scored a median of 41 (interquartile range [IQR], 37.5 to 42.0) of 42 points on the ABILHAND-Kids questionnaire, which assesses manual activity. They also indicated a score of 100% (IQR, 95.6% to 100%)—indicating that they were able to complete all tasks—on the modified PUFI questionnaire. We found a weak correlation between age at the time of follow-up and ABILHAND-Kids score ($r_s = 0.36$, $p < 0.05$), but no significant correlation between age and PUFI score ($r_s = 0.32$, $p > 0.05$). No significant correlations were found between the scores on the various outcome assessment systems and the results on the manual activity questionnaires (r_s ranging from −0.33 to 0.43, $p > 0.1$).

Significant correlation coefficients were found between the examiner-rated VAS scores for function and aesthetics (Fig. 3) and the scores according to the Horii, Cheng, JSSH, Wood, and Dobyns outcome assessment systems. The strongest correlations were between the VAS scores for functional outcome assigned by S.E.R.H. and the scores that S.E.R.H. assigned using the Dobyns and Wood outcome assessment systems ($r_s = 0.93$ and $r_s = 0.86$, respectively). However, the variability among the correlations was considerable for these systems (r_s ranging from 0.33 to 0.93 and 0.34 to 0.86, respectively). Consequently, overall correlations were better for the Horii and JSSH systems (r_s ranging from 0.47 to 0.75 and 0.48 to 0.80, respectively, indicating only moderately strong linear relationships).

Correlation coefficients between the patient-rated VAS scores for function and aesthetics and the outcomes according to the different assessment systems (Fig. 4) were absent to moderate. Overall, however, the strongest significant correlations were for the Horii and JSSH systems (r_s ranging from 0.44 to 0.58 and 0.45 to 0.63, respectively). The average significant correlations between patient-rated VAS scores and the scores according to the outcome assessment systems were $r_s = 0.43$ for function and $r_s = 0.31$ for aesthetics.

Discussion

The aim of this study was to determine which of the available assessment systems is best suited for evaluation of outcome after operative treatment of radial polydactyly. We found that conclusions about overall outcome can vary substantially for a single patient depending on the assessment system used. For example, the results of surgery shown in Figure 5 would be "excellent" if one applies the Wood system criteria, "fair" with use of the JSSH assessment system, and "poor" according to the Tuch system criteria.

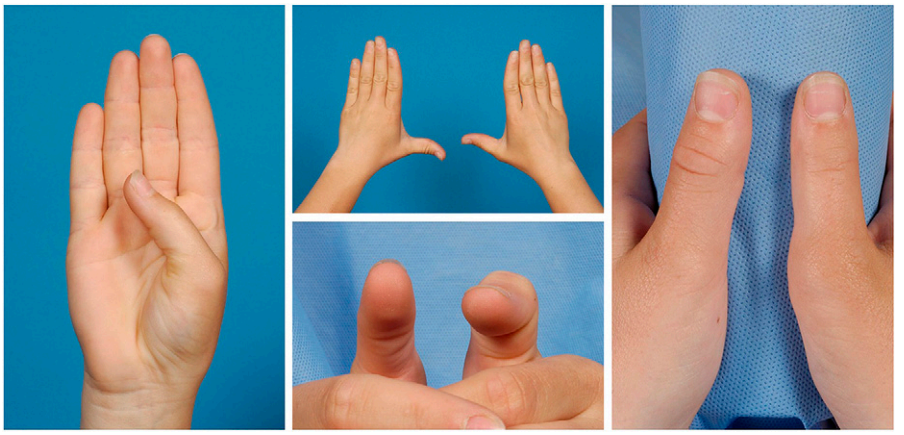

Fig. 5

Postoperative hand function and appearance after treatment of a Flatt type-IV radial polydactyly on the right hand. Slight stability and mobility issues substantially influence the assessment of the overall outcome. A single examiner rated the result for this hand as excellent with the Wood system; good with the Cheng and Dobyns systems; fair with the Tada, Ogino, Larsen, Horii, ALURRA, and JSSH systems; and poor with the Tuch system.

Our results show that the JSSH assessment system has the most optimal properties. It had the highest overall interobserver reliability compared with the other nine assessment systems reviewed. Furthermore, it has relatively high correlations with VAS scores for function and aesthetics. Additionally, observers using the system can differentiate among at least three outcome categories. Finally, it encompasses all important postoperative evaluation parameters. Since we found no correlations between any of the assessment systems and manual activity questionnaires, those results did not contribute to selection of the most optimal system.

Although the Cheng, Tada, and JSSH assessment systems were the three most reliable methods for evaluating outcomes (Table I), the JSSH system has a number of additional assets. First, in contrast to the Cheng and Tada systems, the JSSH system had an ICC that was consistently higher than 0.70. Second, the JSSH assessment system had higher overall correlations with the functional and aesthetic VAS scores. Third, using the Cheng system, two examiners (R.R.D. and C.A.v.N.) could not discriminate between the "poor" and "fair" outcomes, making the system less suitable for detecting improvement after revision surgery. Last, since preoperative patient characteristics influence postoperative residual impairment, use of a solely functional system that emphasizes joint mobility, such as the Tada score, means that a patient with poor preoperative joint motion will never achieve a good outcome according to the system. Moreover, the Cheng and Tada systems do not incorporate residual pain and patient satisfaction as determinants of outcome. We believe that patient satisfaction is an especially important feature.

While the two more subjective systems (Dobyns and Wood) had the highest correlations with the examiner-rated VAS scores, the difference in the correlations for the various examiners using these systems was large. For example, the correlation between the VAS function score and Dobyns system score was 0.93 for one hand surgeon and 0.35 for the other, suggesting that these systems are less reliable.

The absence of significant correlations between any of the systems and the manual activity questionnaires can be explained by the fact that a majority of patients scored "excellent" on these questionnaires (a ceiling effect). From our data, it cannot be determined whether this resulted from radial polydactyly being a relatively mild functional impairment regardless of the success of the surgery or it was due to these questionnaires not being sensitive enough to discriminate between good and poor results of the treatment of thumb-related problems.

The main limitations of this study are the relatively small sample size and the fact that all three examiners did not evaluate every patient. On the basis of the originally described frequency4 of Flatt type-II and IV radial polydactyly and the total number of our patients (211), one would expect approximately 122 patients to have type II or IV. The main criteria limiting inclusion to fifty-eight patients were an age of four years or older, a minimum of one year of postoperative follow-up, and the relatively high occurrence of radial polydactyly with triphalangeal components in our population[30]. Nevertheless, we were able to show considerable differences in reliability, discriminative properties, and correlation consistency among the different systems. This was the aim of the study, and we are confident that our sample size was adequate for that purpose. Increasing sample size might influence correlations with questionnaires that failed to reach the significance level of $p < 0.05$ in the current setting. However, comparable studies have also shown poor correlations between outcomes involving body structures and function and the scores on questionnaires on manual activity or impairment[24,25].

Another limitation is the validity of the ABILHAND-Kids and PUFI questionnaires as measures of manual ability, as these questionnaires were validated only for use for children with cerebral palsy[27] and populations with transverse upper-limb deficiencies[31]. Although there are alternative measures of manual activity, such as the Assisting Hand Assessment (AHA)[32], most of them are very time-consuming and require specific expertise of hand therapists. In contrast, filling out a questionnaire and completing the JSSH assessment system form takes, in total, about fifteen minutes, which is more feasible in a busy outpatient clinic.

A possible third limitation of our study is the restriction of our study population to the most prevalent types of radial polydactyly—types II and IV. Although this might impair the generalizability of the results, Flatt types II and IV affect interphalangeal and metacarpophalangeal joints, making evaluation of alignment, stability, mobility, and strength more relevant. These items are incorporated in most outcome assessment systems.

One of the strengths of this study was the combined assessment of patients by two surgeons highly experienced with treating congenital hand differences and an independent examiner in the analysis of interobserver reliability. The independent examiner was qualified to conduct patient evaluation but lacked the surgical experience of the two frequently collaborating hand surgeons. This is a realistic research setting because clinical research is often conducted either by observers who are in some stage of medical training or by physician assistants, due to the busy schedules of experienced consultants. Indeed, we did find considerable differences in interobserver reliability between the two hand surgeons and the independent examiner for the majority of the assessment systems. Although the surgical inexperience of the independent examiner may have lowered the overall agreement, we believe that the validity and generalizability of an assessment system is increased when it can be reliably applied irrespective of surgical experience in the field of congenital hand differences. Another strength of the study was that we related the patients' satisfaction with the function and aesthetic results to the outcome assessments by the three examiners. Our findings suggest poor-to-moderate correlations between physician-rated and patient-rated outcome, which is well known[2,6,8,24,33]. Moreover, patient satisfaction is an important reason for seeking medical attention, a determining factor for performing revision surgery, and a substantial confounder in outcome evaluation after revision surgery. We therefore believe that patient satisfaction should be an integral part of outcome assessment, as it is in the JSSH system.

A third strength of the study was the systematic evaluation of the validity of the outcome assessment systems. Validity evaluation usually includes assessment of content validity (i.e., whether the system measures all facets of the construct that it is supposed to measure, which, in the present study, was the overall outcome of the treatment of the radial polydactyly), construct validity (i.e., whether the system measures or correlates with the construct that it claims to measure), and criterion validity (i.e., the extent to which results obtained with the system relate to concrete criteria in the "real world"). In the absence of a gold standard to describe hand function as an outcome of the treatment of radial polydactyly, we assessed content validity on the basis of consensus, construct validity on the basis of correlations with VAS scores, and criterion validity on the basis of correlation with manual activity questionnaires.

A general limitation of the assessment systems is their item weighting. For example, the JSSH system assigns the same number of points to metacarpophalangeal joint stability as it does to interphalangeal joint flexion. Currently, these item weights are not based on the items' effect on overall hand function. Possibly, the JSSH system could benefit from a redistribution of points over the various items, based on their impact on overall hand function.

In this study, we compared the reliability and validity of ten assessment systems designed to evaluate outcome and residual impairment in patients with radial polydactyly. The Tada score (or modified versions of this assessment system) has been applied most frequently to describe outcomes of treatment of radial polydactyly[7-16]. However, the results of our study indicate that the JSSH system is better suited for this purpose. In addition to better reliability and overall correlations, the JSSH encompasses more items that we think are relevant to describe overall outcome. We therefore recommend the JSSH assessment system for future scientific evaluation of outcome and residual impairment after surgery for radial polydactyly.

Appendix

eA A table comparing the characteristics of the study population with those of the eligible patients who did not participate in the study is available with the online version of this article as a data supplement at jbjs.org. ■

Robert R. Dijkman, MD
Ruud W. Selles, PhD
Department of Plastic and Reconstructive Surgery and Hand Surgery,
Erasmus MC, P.O. Box 2040,
Room EE.15.91,
3000 CA Rotterdam, the Netherlands.
E-mail address for R.R. Dijkman: r.dijkman@erasmusmc.nl.
E-mail address for R.W. Selles: r.selles@erasmusmc.nl

Christianne A. van Nieuwenhoven, MD, PhD
Department of Plastic and Reconstructive Surgery and Hand Surgery,
Erasmus MC, P.O. Box 2040,
Room HS 511,
3000 CA Rotterdam, the Netherlands.
E-mail address: c.vannieuwenhoven@erasmusmc.nl

Steven E.R. Hovius, MD, PhD
Department of Plastic and Reconstructive Surgery and Hand Surgery,
Erasmus MC, P.O. Box 2040,
Room HS 501,
3000 CA Rotterdam, the Netherlands.
E-mail address: s.e.r.hovius@erasmusmc.nl

References

1. Ekblom AG, Laurell T, Arner M. Epidemiology of congenital upper limb anomalies in 562 children born in 1997 to 2007: a total population study from stockholm, sweden. J Hand Surg Am. 2010 Nov;35(11):1742-54. Epub 2010 Oct 20.

2. Goldfarb CA, Patterson JM, Maender A, Manske PR. Thumb size and appearance following reconstruction of radial polydactyly. J Hand Surg Am. 2008 Oct;33(8):1348-53.

3. Takagi R, Kawabata H, Matsui Y. Thumb polydactyly with symphalangism in young children. J Hand Surg Eur Vol. 2009 Dec;34(6):800-4. Epub 2009 Jul 10.

4. Wassel HD. The results of surgery for polydactyly of the thumb. A review. Clin Orthop Relat Res. 1969 May-Jun;64(64):175-93.

5. Dobyns JH, Lipscomb PR, Cooney WP. Management of thumb duplication. Clin Orthop Relat Res. 1985 May;195(195):26-44.

6. Tuch BA, Lipp EB, Larsen IJ, Gordon LH. A review of supernumerary thumb and its surgical management. Clin Orthop Relat Res. 1977 Jun;(125):159-67.

7. Tada K, Yonenobu K, Tsuyuguchi Y, Kawai H, Egawa T. Duplication of the thumb. A retrospective review of two hundred and thirty-seven cases. J Bone Joint Surg Am. 1983 Jun;65(5):584-98.

8. Kemnitz S, De SL. Pre-axial polydactyly: outcome of the surgical treatment. J Pediatr Orthop B. 2002 Jan;11(1):79-84.

9. Yen CH, Chan WL, Leung HB, Mak KH. Thumb polydactyly: clinical outcome after reconstruction. J Orthop Surg (Hong Kong). 2006 Dec;14(3):295-302.

10. Maillet M, Fron D, Martinot-Duquennoy V, Herbaux B. Results after surgical treatment of thumb duplication: a retrospective review of 33 thumbs. J Child Orthop. 2007 Jul;1(2):135-41. Epub 2007 May 05.

11. Tonkin MA, Bulstrode NW. The Bilhaut-Cloquet procedure for Wassel types III, IV and VII thumb duplication. J Hand Surg Eur Vol. 2007 Dec;32(6):684-93. Epub 2007 Sep 25.

12. Iwasawa M, Noguchi M, Mishima Y, Fujita K. Long-term results of nail fusion plasty of the duplicated thumb. J Plast Reconstr Aesthet Surg. 2008 Sep;61(9): 1085-9. Epub 2007 Sep 21.

13. Ogino T, Ishii S, Takahata S, Kato H. Long-term results of surgical treatment of thumb polydactyly. J Hand Surg Am. 1996 May;21(3):478-86.

14. Horii E, Nakamura R, Sakuma M, Miura T. Duplicated thumb bifurcation at the metacarpophalangeal joint level: factors affecting surgical outcome. J Hand Surg Am. 1997 Jul;22(4):671-9.

15. Larsen M, Nicolai JP. Long-term follow-up of surgical treatment for thumb duplication. J Hand Surg [Br]. 2005 Jun;30(3):276-81. Epub 2005 Apr 07.

16. Tien YC, Chih TT, Wang TL, Fu YC, Chen JC. Soft tissue reconstruction for type IV-D duplicated thumb: a new surgical technique. J Pediatr Orthop. 2007 Jun;27(4): 462-6.

17. Cheng JC, Chan KM, Ma GF, Leung PC. Polydactyly of the thumb: a surgical plan based on ninety-five cases. J Hand Surg Am. 1984 Mar;9(2):155-64.

18. Brouwer N, van der Biezen JJ. After thumb duplication: a proposal for evaluation of the result. The ALURRA system. Eur J Plast Surg. 2006 Apr;29:1-4. Epub 2006 Apr 25.

19. JSSH Congenital Hand Committee. Evaluation sheet for polydactyly of the thumb [in Japanese]. J Japan Soc Surg Hand. 2007;24:422.

20. Flatt AE. The care of congenital hand anomalies. 2nd ed. St. Louis: Quality Medical Publishers; 1994.

21. Wood VE. Congenital hand deformities. In: Green DP, editor. Operative hand surgery. 2nd ed. New York: Churchill Livingstone; 1988. p 447-58.

22. 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 Jul-Sep;22(3):271-6, quiz :277. Epub 2009 Mar 10.

23. Kapandji A. [Clinical test of apposition and counter-apposition of the thumb]. Ann Chir Main. 1986;5(1):67-73.

24. Goldfarb CA, Klepps SJ, Dailey LA, Manske PR. Functional outcome after centralization for radius dysplasia. J Hand Surg Am. 2002 Jan;27(1):118-24.

25. Zuidam JM, de Kraker M, Selles RW, Hovius SE. Evaluation of function and appearance of adults with untreated triphalangeal thumbs. J Hand Surg Am. 2010 Jul;35(7):1146-52. Epub 2010 Jun 19.

26. Bailey B, Daoust R, Doyon-Trottier E, Dauphin-Pierre S, Gravel J. Validation and properties of the verbal numeric scale in children with acute pain. Pain. 2010 May;149(2):216-21. Epub 2010 Feb 25.

27. Arnould C, Penta M, Renders A, Thonnard JL. ABILHAND-Kids: a measure of manual ability in children with cerebral palsy. Neurology. 2004 Sep 28;63(6): 1045-52.

28. Wright FV, Hubbard S, Jutai J, Naumann S. The Prosthetic Upper Extremity Functional Index: development and reliability testing of a new functional status questionnaire for children who use upper extremity prostheses. J Hand Ther. 2001 Apr-Jun;14(2):91-104.

29. Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. Psychol Bull. 1979 Mar;86(2):420-8.

30. Zuidam JM, Ananta M, Hovius SER. Triplicated thumbs: a rarity? J Plast Reconstr Aesthet Surg. 2008 Sep;61(9):1078-84. Epub 2007 Sep 05.

31. Wright FV, Hubbard S, Naumann S, Jutai J. Evaluation of the validity of the prosthetic upper extremity functional index for children. Arch Phys Med Rehabil. 2003 Apr;84(4):518-27.

32. Krumlinde-Sundholm L, Eliasson AC. Development of the Assisting Hand Assessment: a Rasch-built intended for children with unilateral upper limb impairments. Scand J Occup Ther. 2003;10:16-26.

33. Andrew JG, Sykes PJ. Duplicate thumbs: a survey of results in twenty patients. J Hand Surg [Br]. 1988 Feb;13(1):50-3.