Smartphone-based Measurement of Post-traumatic Elbow Range of Motion: Agreement with Goniometry and Clinical Assessment

Dror Ronel MD[1], Galina Cohen Shapiro MD PhD[1,2,3,4], Nadav Rinott MD[1,5], Avi Fishbein MD[1], and Yaniv Keren MD[1]

[1]Department of Orthopedics, Rambam Health Care Campus, Haifa, Israel
[2]Medical Corps, Surgeon General's Headquarters, Israel Defense Forces, Ramat Gan, Israel
[3]Department of Military Medicine and "Tzameret", Faculty of Medicine, Hebrew University of Jerusalem, Jerusalem, Israel
[4]Institute for Research in Military Medicine (IRMM), Faculty of Medicine, Hebrew University of Jerusalem, Jerusalem, Israel
[5]Department of Pediatric Orthopedics, Rappaport Faculty of Medicine, Technion–Israel Institute of Technology, Haifa, Israel

ABSTRACT

Background: Elbow range of motion (ROM) is commonly assessed during initial evaluation and follow-up. While goniometry is considered the current gold standard, previous studies have raised concerns regarding its user dependence and measurement consistency.

Objective: To evaluate whether smartphone application-based elbow ROM measurements are comparable to visual estimation and goniometry in patients following elbow trauma.

Methods: We conducted a prospective cohort study of 18 patients with elbow injuries. Demographic and medical data were collected. Maximal flexion-extension active ROM was measured by goniometry, smartphone application, and visual estimation by orthopedic surgeons. Measurements were compared to determine agreement and accuracy among different modalities.

Results: Eighteen patients with elbow injuries were included. Inter-rater reliability was good to excellent across all rater groups (intraclass correlation coefficient 0.78–0.93). Using paired *t*-tests with Bonferroni correction, only one attending surgeon showed a significant underestimation of extension compared to goniometry (16.9° vs. 22.5°, $P < 0.05$). No significant differences were found for residents, other attendings, or smartphone measurements in either extension or flexion ($P > 0.05$). Bland-Altman analysis revealed modest mean bias within ± 6° for all methods, although limits of agreement were wide (up to ± 20–30°). Smartphone-based measurements demonstrated significantly higher and less variable Pearson correlation coefficients with goniometry compared to surgeon visual estimations (0.91–0.95 vs. 0.81–0.93, $P < 0.05$).

Conclusions: Smartphone-based elbow ROM measurements were statistically comparable to goniometry and demonstrated superior correlation and accuracy relative to clinical visual assessments. These findings support the integration of smartphone tools in the clinical setting.

IMAJ 2026; 28: 590–594

KEY WORDS: elbow trauma, goniometry, range of motion (ROM), rehabilitation, smartphone-based measurements

Elbow range of motion (ROM) assessment is an important part of a physical examination as well as for clinical follow-up. It is commonly used in assessing postoperative outcomes, and monitoring rehabilitation and is regarded as an objective outcome measurement. The most common instrument used for joint ROM measurement is the universal goniometer because it is easy to use, affordable, and reliable [1,2]. Nonetheless, in hinge-type joints it may yield inconsistent measurements as the joint is difficult to mark and stabilize [3]. User inexperience may increase measurement error rates [4,5]. Compared with visual ROM estimation, it has been suggested that a goniometer should be used when seeking accuracy of up to 10° [6].

Smartphone applications based on photography [7-10] or an accelerometer [11-14], and more recently machine-learning based approaches, offer a potentially highly reliable and easily accessible approach to ROM measurement. Additional potential benefits include ease of use and a quicker measurement [15]. Most studies thus far have demonstrated adequate intra-rater or inter-rater reliability and/or validity for smartphone application-based ROM measurement overall [3].

For elbow ROM measurements specifically, very few studies have been published to date [3,15,16], so it is unclear whether smartphone application-based measurement is a valid and reliable measurement tool. Smartphone-based measurement has demonstrated high validity and reliability in controlled settings [7,16-18]. The elbow joint is prone to post-traumatic stiffness, and patient satisfaction has been associated with ROM outcomes, suggesting accurate ROM measurements in this patient population is clinically important [19,20].

This study adds to the existing literature by evaluating smartphone-based elbow ROM measurement in a real-world orthopedic trauma population, directly comparing goniometry and clinical visual assessments. We hypothesized that a smartphone application-based elbow ROM measurement would be highly correlated with visual estimation and goniometry in patients after elbow trauma.

PATIENTS AND METHODS

This prospective cohort study included 18 patients who presented to an orthopedic trauma outpatient clinic with an elbow injury, including either a distal humerus fracture or a proximal radius or proximal ulna fracture during 2020–2021. A post-hoc power analysis was performed for the primary comparisons. A post-hoc power calculation for detecting a mean paired difference of 5 degrees, assuming a standard deviation of paired differences of 5 degrees, alpha 0.05, and n=18, yielded a power of 97.9%. The study was approved by the institutional review board (approval number 0263-20). Patients were included in the surgical care group if they had undergone surgery before their first outpatient follow-up appointment.

DATA COLLECTION AND STUDY VARIABLES

Age, sex, diagnosis, and treatment type (conservative or surgical) were collected and analyzed from the electronic health records. Maximal flexion-extension active ROM was measured for each patient using a goniometer and by a smartphone application. The measurements were all conducted on the same smartphone (Apple iPhone 11) using Orthophysical version 1.3 (by Software Marketing Bank). Measurements were conducted with the app using both photographic analysis and accelerometer modes. All measurements were conducted by the same orthopedic surgeon in the same outpatient clinic setting. Active ROM was measured with the camera at a distance of 1–1.5 meters from the patient. Full extension of the elbow was defined as 0 degrees. This app was chosen because it is free, simple, and allows both photographic and accelerometer measurements. In addition, a visual estimation was recorded by two orthopedic surgery residents and two attending orthopedic surgeons, all blinded to previous measurements.

Patients were examined for active flexion and extension, standing when possible. There was one trial per measurement, and all the measurements were conducted at the same visit. The residents were senior residents who completed at least 3 years of the residency and successfully passed the first part of the Israeli orthopedic surgery residency exam. Attending orthopedic surgeons were all Israeli board qualified orthopedic surgeons. Visual estimation was evaluated independently so each physician was unaware of the estimations of other physicians, goniometers, or smartphone application measurements.

Patient-reported outcomes were not available in our study population. Even though the sample had sufficient power, it was not large enough to add subgroup analysis.

Table 1. Patient characteristics

Patients	(n=18)
Age at injury (years), mean ± standard deviation	48 ± 16
Sex, n (%)	
Male	11 (61%)
Female	7 (39%)
Fracture	
Humerus	12 (67%)
Radius	3 (17%)
Ulna	4 (22%)
Treatment	
Conservative	11 (69%)
Surgical	5 (31%)

STATISTICAL ANALYSIS

Paired *t*-tests with Bonferroni correction for multiple comparisons were used to compare measured ranges of motion to goniometry. Inter-observer agreement was assessed using the interclass correlation coefficient (ICC), applying a two-way random-effects model with absolute agreement (ICC(2,1) for single measurements and ICC(2,k) for averaged measurements). Agreement between measurement methods and goniometry was evaluated using Bland–Altman analysis, calculating mean bias, and 95% limits of agreement (mean difference ± 1.96 standard deviations). Plots were visually inspected for systematic bias and heteroscedasticity. Separate analyses were performed for flexion and extension. Pearson's correlation coefficients were calculated between measured ranges of motion and goniometry. P-value < 0.05 was considered statistically significant. The statistical analysis was performed using Microsoft Excel™ 2021 Version 17 (Microsoft® Corporation, Redmond, WA, USA) and R Statistical Software, version 4.4.3 (R Foundation for Statistical Computing, Vienna, Austria).

RESULTS

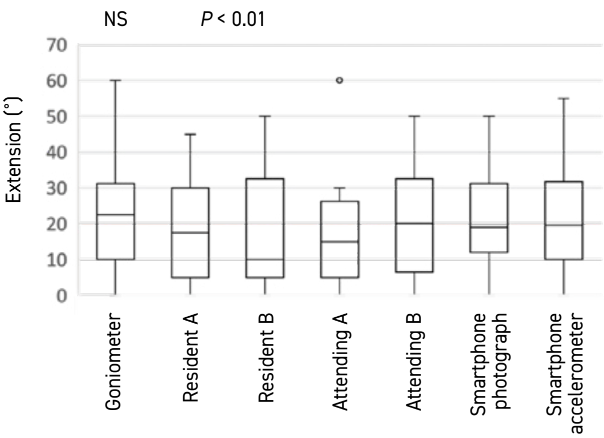
Eighteen patients with significant elbow injuries were included in this study [Table 1]; 11 males and 7 females with an average age of 48 years. Most had presented a distal humerus fracture, while some had either a proximal radius or proximal ulna fracture. One patient had a proximal ulna fracture and a distal radius fracture. Most patients were treated conservatively.

The interclass correlation coefficient (ICC) between resident raters was 0.78 (95% confidence interval [95%CI] 0.51–0.91), indicating good inter-rater reliability. For senior raters, the ICC was 0.84 (95%CI 0.63–0.94), also demonstrating excellent reliability. Averaging ratings further improved reliability to 0.88 for residents and 0.91 for seniors. The ICC for all raters combined was 0.78 (95%CI 0.61–0.90) for single measurements. When averaging ratings across raters, reliability increased further to 0.93 (95%CI 0.86–0.97).

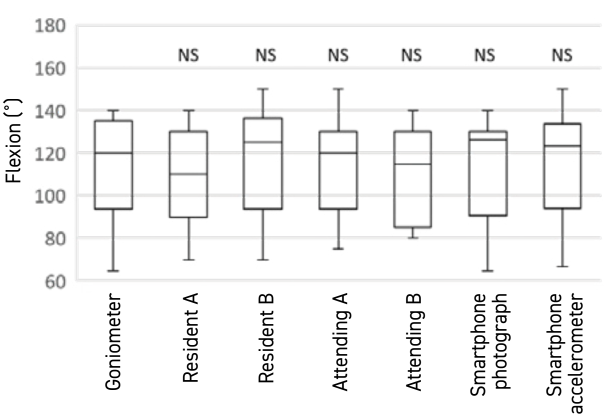
Figure 1. Measured by goniometry, orthopedic surgeons and a smartphone application

NS = not significant
Circle denotes an outlier

[A] Extension



[B] Flexion



When ROMs in extension were compared to those measured using goniometry, using paired *t*-test with Bonferroni correction, only one senior orthopedic surgeon had underestimated extension as demonstrated by lower average extension angle 16.9˚ (95%CI 13–20) compared to an average of 22.5˚ (95%CI 19–26) measured by goniometry. No significant differences were found when comparing extension by both residents, other senior surgeons, smartphone accelerometer, or photograph analysis to goniometry ($P > 0.05$) [Figure 1]. No significant differences were found when comparing ROMs in flexion as measured by goniometry to any orthopedic surgeon or smartphone applications ($P > 0.05$) [Figure 1].

Bland-Altman analysis demonstrated that, for both extension and flexion, all alternative methods (including resident and attending raters as well as smartphone-based measurements) showed only modest mean bias when compared to goniometry, typically within ± 6˚. For extension, all methods exhibited a small positive bias, indicating a tendency to overestimate extension angles relative to the reference standard. However, across both movements, the limits of agreement were relatively wide, ranging up to approximately 20 to 30˚ in some cases. This finding suggests that, while group averages are similar, substantial variability exists at the individual level. Notably, the smartphone camera method showed the smallest bias and narrower limits of agreement compared to other methods in extension but still permitted differences up to 13 to 14˚.

Very strong correlations were found between measurements by orthopedic surgeons and a smartphone application and goniometry with Pearson's correlation coefficients between 0.813 to 0.948 overall [Table 2].

Figure 2. Correlation coefficients calculated between ranges of motion as measured by orthopedic surgeons and a smartphone application

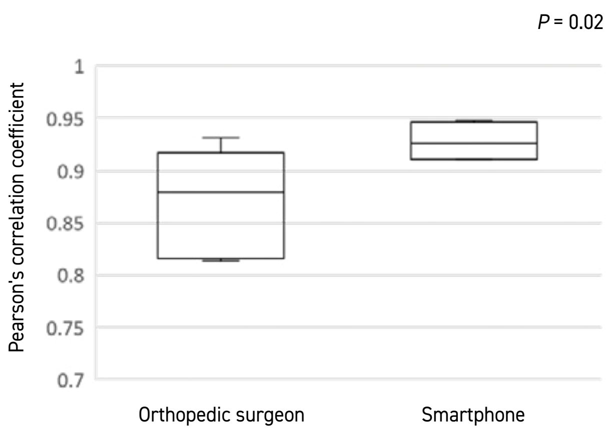


Table 2. Correlation coefficients calculated between ranges of motion as measured by orthopedic surgeons and a smartphone application

	Resident A	Resident B	Attending A	Attending B	Smartphone photograph	Smartphone accelerometer
Extension	0.911	0.860	0.899	0.813	0.911	0.911
Flexion	0.931	0.815	0.919	0.819	0.940	0.948

A statistically significant difference was found between the correlation coefficients of measurements using a smartphone (accelerometer and photograph analysis combined) compared to the correlation coefficients of measurements of orthopedic surgeons ($P < 0.05$) [Figure 2]. Importantly, the correlation coefficients of smartphone-based measurements appear to be higher and less variable.

DISCUSSION

We compared elbow ranges of motion in 18 patients after elbow trauma measured by goniometry, orthopedic surgeons, and a smartphone application. Goniometry is the commonly accepted method and was used as the reference for statistical comparison [1,2]. Our results indicate that smartphone-based measurements were not statistically different from those of goniometry. Furthermore, smartphone-based measurements were found to have better averages and had lower limits of agreement, supporting the use of smartphone-based measurements over clinical evaluations.

Previous studies have demonstrated systematic differences between measurement methods, with smartphone applications tending to overestimate flexion while clinical estimation may underestimate it [14]. Like our findings, agreement at the group level did not translate into perfect agreement at the individual level.

Our findings are consistent with previous studies demonstrating excellent agreement between smartphone applications and goniometry, and superior performance compared to visual estimation [12]. However, in contrast to prior reports with narrower limits of agreement [13], our study demonstrated wider variability, likely reflecting a more heterogeneous post-traumatic adult population. Measurement variability is known to increase when transitioning from standardized angles to real human joint motion, even with highly reliable devices [13]. This result may explain the wider limits of agreement observed in our cohort.

Smartphone-based elbow ROM measurements demonstrated comparable results to goniometry at the group level. However, the relatively wide limits of agreement indicated substantial variability at the individual level. Therefore, smartphone-based assessment may serve as a useful adjunct tool but cannot be considered a replacement for goniometric measurement.

While previous studies [6] showed that clinical evaluation is accurate with more experienced physicians, our study found mixed results with no clear advantage to measurements conducted by attendings compared to residents. One of the factors that could have contributed to this finding is the fact that the residents who participated in this study were senior residents, with a minimum of post-graduation of year 4. Likewise, even for statistically significant differences between goniometry and clinical estimation, the average difference in measurement was between 5˚ and 12.5˚, which is probably at the lower end of what is clinically significant.

Previous studies have put the emphasis on self-measurement or measurement by layperson [7,15]. Our study focused on clinical use by orthopedic surgeons at a trauma clinic, applying it to a real-world cohort of trauma patients. The findings suggest that digital tools may offer a reliable, objective alternative for ROM assessment in orthopedic trauma patients, challenging assumptions about the superiority of clinical experience and supporting broader integration of mobile technology into routine musculoskeletal evaluation.

This study has several limitations. First, the sample size was relatively small and derived from a single center, which may limit generalizability. Second, measurements were obtained in a single session without repeated trials, precluding assessment of intra-rater reliability. Third, only flexion and extension were measured, thus lacking information about forearm rotation. In addition, patient-reported outcomes were not included, limiting assessment of clinical relevance. Last, the inherent variability of post-traumatic elbow conditions may have contributed to the wide limits of agreement observed. While the wide limits of agreement represent a limitation, they likely reflect real-world clinical variability and may provide a more realistic estimate of measurement performance in routine practice.

CONCLUSIONS

We showed equivalent and sometimes slightly better results while using the orthophysical app over clinical assessment.

Even though these findings only relate to the specific app tested, they support the use of smartphone-based measurements. Smartphone-based measurements demonstrated mean bias and limits of agreement that were comparable to or narrower than those of human raters, supporting their potential use. As technology becomes more integrated in clinical care there is an expectation for improving physical examination accuracy and reliability and reducing provider-experience-related errors by using applications for better ROM measurements. Possible obstacles for integrating these technologies into routine clinical practice may include information security concerns, a current lack of cost-benefit analysis, and a lack of standardization between applications and health-care providers.

Building on prior studies demonstrating validity in controlled or non-traumatic populations, in this study, we evaluated the agreement between measurement methods in a real-world orthopedic trauma setting. The results add to the existing literature by addressing the limited evidence on measurement agreement under routine clinical conditions. Smartphone-based measurements are not yet intended to replace the goniometer as the gold standard, but they are a good adjunct that adds reliability compared to clinical visual estimation.

Correspondence

Dr. D. Ronel
Dept. of Orthopedics, Rambam Health Care Campus, Haifa 3109601, Israel
Email: dror.ronel@gmail.com

References

1. van Rijn SF, Zwerus EL, Koenraadt KL, Jacobs WC, van den Bekerom MP, Eygendaal D. The reliability and validity of goniometric elbow measurements in adults: a systematic review of the literature. *Shoulder Elb* 2018; 10 (4): 274-84.
2. Furness J, Schram B, Cox AJ, Anderson SL, Keogh J. Reliability and concurrent validity of the iPhone(®) Compass application to measure thoracic rotation range of motion (ROM) in healthy participants. *PeerJ* 2018; 6: e4431.
3. Keogh JWL, Cox A, Anderson S, et al. Reliability and validity of clinically accessible smartphone applications to measure joint range of motion: a systematic review. *PLoS One* 2019; 14 (5): e0215806.
4. Cronin J, Nash M, Whatman C. Assessing dynamic knee joint range of motion using siliconcoach. *Phys Ther Sport Off J Assoc Chart Physiother Sport Med* 2006; 7 (4): 191-4.
5. Piriyaprasarth P, Morris ME, Winter A, Bialocerkowski AE. The reliability of knee joint position testing using electrogoniometry. *BMC Musculoskelet Disord* 2008; 9: 6.
6. Luria S, Apt E, Kandel L, Bdolah-Abram T, Zinger G. Visual estimation of pro-supination angle is superior to wrist or elbow angles. *Phys Sportsmed* 2015; 43 (2): 155-60.
7. Meislin MA, Wagner ER, Shin AY. A Comparison of elbow range of motion measurements: smartphone-based digital photography versus goniometric measurements. *J Hand Surg Am* 2016; 41 (4): 510-515.e1.
8. Martins I, Gunot M, Dias A. Validity and reliability of a smartphone application versus 2d software for joint range of motion measurement: a cross-sectional validation study. *Muscles* 2025; 4 (1): 8.
9. Meng Y, Lee M, Clark H-JY, et al. Evaluation of range of motion using vision artificial intelligence (AI) in musculoskeletal medicine. *J Med Artif Intell* 2024; 7: 22.
10. Kim L, Kim H, Seungjun K, Lee J, Cho J, Chang E. Reliability and agreement of a smartphone-based markerless analysis system (Sports2D) for joint range of motion measurement. *Korean J Sport Med* 2025; 43: 254-64.
11. Pourahmadi MR, Ebrahimi Takamjani I, Sarrafzadeh J, et al. Reliability and concurrent validity of a new iPhone(®) goniometric application for measuring active wrist range of motion: a cross-sectional study in asymptomatic subjects. *J Anat* 2017; 230 (3): 484-95.
12. Koong DP, Lee J, Cheng TL, Little DG. Validity and reliability of smartphone inclinometer applications for measurement of elbow range of motion in paediatric patients. *J Child Orthop* 2020; 14 (5): 488-94.
13. Kiatkulanusorn S, Luangpon N, Srijunto W, et al. Analysis of the concurrent validity and reliability of five common clinical goniometric devices. *Sci Rep* 2023; 13 (1): 20931.
14. Vauclair F, Aljurayyan A, Abduljabbar FH, et al. The smartphone inclinometer: a new tool to determine elbow range of motion? *Eur J Orthop Surg Traumatol* 2018; 28 (3): 415-21.
15. Keijsers R, Zwerus EL, van Lith DRM, et al. Validity and reliability of elbow range of motion measurements using digital photographs, movies, and a goniometry smartphone application. *J Sport Med (Hindawi Publ Corp)* 2018; 2018: 7906875.
16. Maheshkumar P.G., Harihara Subramanyan, Swapna N. JJM. A study to analyse the reliability and validity of the smartphone goniometer application (G-PRO) by measuring the elbow joint range of motion. *Indian J Public Heal Res Dev* 2020; 11 (3 SE-Articles): 235-40.
17. Behnoush B, Tavakoli N, Bazmi E, et al. Smartphone and universal goniometer for measurement of elbow joint motions: a comparative study. *Asian J Sports Med* 2016; 7 (2): e30668.
18. Lin NCJ, Hayward KS, D'Cruz K, Thompson E, Li X, Lannin NA. Validity and reliability of a smartphone inclinometer app for measuring passive upper limb range of motion in a stroke population. *Disabil Rehabil* 2020; 42 (22): 3243-9.
19. Morrey BF, Askew LJ, Chao EY. A biomechanical study of normal functional elbow motion. *J Bone Joint Surg Am* 1981; 63 (6): 872-7.
20. Sardelli M, Tashjian RZ, MacWilliams BA. Functional elbow range of motion for contemporary tasks. *J Bone Joint Surg Am* 2011; 93 (5): 471-7.

Our greatest strength is the power of our example, not just the example of our power.
Joseph Robinette Biden Jr. (Born 1942), American politician, 46th president of the United States

An artist should never be a prisoner of himself, prisoner of style, prisoner of reputation, prisoner of success, etc.
Henri Matisse (1869–1954), French visual artist, known for both his use of color and his fluid and original draftsmanship