J Shoulder Elbow Surg (2026) 35, 517–525

ELSEVIER

Journal of Shoulder and Elbow Surgery

www.elsevier.com/locate/ymse

Dynamic track-based modification (Dynamic Track Instability Management Score) of current shoulder instability scores is a highly reliable alternative for decision making in anterior instability

Deepak N. Bhatia, MS(Orth), DNB(Orth)[a,b,*], Narottam Das, MS(Orth)[a], Parimal Malviya, MS(Orth)[a]

[a]*SportsDocs, Mumbai, Maharashtra, India*
[b]*Department of Orthopaedic Surgery, Sir HN Reliance Foundation Hospital and Research Center, Mumbai, Maharashtra, India*

Background: Shoulder instability scoring systems (Instability Severity Index [ISI] Score and the Glenoid Track Instability Management Score [GTIMS]) use radiographic or computed tomographic (CT) scan measurements for scoring bone loss. Radiologic methods are known to have variable reliability, and this variability can affect decision making. The purpose of this study was to assess reliability, reproducibility, and diagnostic validity of a new dynamic track-based score, the Dynamic Track Instability Management Score (DTIMS), and to compare these with current gold standard scores for predicting the necessity for a bony reconstruction procedure. The hypothesis was that DTIMS will have greater agreement and inter-observer reliability (IOR) for bone defect assessment, and in decision making for soft tissue vs. bony surgery.

Methods: Between February 2023 and October 2024, a total of 108 patients with traumatic anterior shoulder instability were evaluated clinically and using radiographs and CT scans, by 2 independent observers. Radiographs were used to identify glenoid and humeral head defects, and 3D CT scans were used to classify defects as on- and off-track. The data were used to calculate 2 shoulder scores (GTIMS and ISI score). During arthroscopy, a Dynamic Arthroscopic Standardized Test (DAST) was used to classify defects into on- and off-track, and DTIMS was calculated. Cohen kappa was used to evaluate IOR for bone loss assessment for the ISI score (Hill-Sachs on external rotation and loss of inferior glenoid contour on anteroposterior radiograph), and for DTIMS and GTIMS (on- and off-track), and for determining the need for bone or soft tissue surgery for all 3 scores. IOR for total scores was assessed using the intraclass correlation coefficient (ICC). Diagnostic validity of DTIMS was assessed against the ISI score and GTIMS, and sensitivity, specificity, positive and negative predictive values, and diagnostic accuracy were reported.

Results: IOR for total score was excellent (ICC >0.9) for all 3 scores. IOR for identification of significant bone defects was near perfect ($\kappa = 0.94$, $P < .001$) in DTIMS, and was substantial ($\kappa = 0.69$, $P < .001$) in GTIMS. However, IOR was moderate ($\kappa = 0.47$ glenoid, 0.58 Hill-Sachs, $P < .001$) with the radiographic method used in ISI score. Similarly, IOR for surgical decision making was near perfect for DTIMS ($\kappa = 0.91$, $P < .001$) and GTIMS ($\kappa = 0.85$, $P < .001$) and was lower ($\kappa = 0.74$, $P < .001$) for ISI score. DTIMS showed the highest sensitivity (89%, 91%) for predicting a bony procedure, and ISI score showed the highest specificity (78%, 81%) when compared with GTIMS.

The Institutional Ethics Committee of Sir HN Reliance Foundation Hospital and Research Center, Mumbai, India, has approved this study (protocol no. HNH/IEC/2024/OCS/ORTH/150).

*Reprint requests: Deepak N. Bhatia, MS(Orth), DNB(Orth), SportsDocs, 4B Summerville, Bandra West, Mumbai, Maharashtra 400050, India.
E-mail address: sportsdocs50@gmail.com (D.N. Bhatia).

https://doi.org/10.1016/j.jse.2025.05.031

Conclusion: Substituting the image-based assessment of bone defects with dynamic track method (ie, DAST) resulted in a highly reliable scoring system (DTIMS) that can be used as an alternative to current gold standard scores (ISI score and GTIMS) for decision-making in anterior shoulder instability.

Level of evidence: Basic Science Study; Development and Validation of Classification System

Keywords: Shoulder; instability; bone defect; glenoid track; Hill-Sachs lesion; arthroscopy; shoulder score; off-track

Decision making in shoulder instability is based on several clinical factors and on imaging-based assessment of glenohumeral bone loss. Shoulder instability scoring systems use a summation of risk factors to identify failure risk after an arthroscopic soft tissue repair. Two commonly used scores are the Instability Severity Index (ISI) score and the Glenoid Track Instability Management Score (GTIMS).[2,12] Both scores are based on a 10-point scale in which 6 points are determined by clinical factors (age <20 years [2 points], involvement in competitive sports [2 points], contact sports or with forced overhead activity [1 point], and shoulder hyperlaxity [1 point]) and 4 points are allocated to significant bone defects. The ISI score score uses radiographic criteria to recognize a significant Hill-Sachs lesion (HSL; anteroposterior shoulder radiograph in external rotation [2 points]) and a significant glenoid defect (inferior glenoid contour loss on a true anteroposterior shoulder radiograph [2 points]). Di Giacomo et al[12] proposed a modification of the ISI score that replaced the radiographic determination of bone defect severity in ISI score with a computed tomographic (CT) scan–based on- and off-track classification and allocated 4 points to an off-track lesion in GTIMS. Further studies have reported on the disadvantages of the CT scan method: (1) CT or magnetic resonance imaging–based measurements are reported to have poor and variable reliability[14,24]; (2) when measured under dynamic conditions, glenoid track widths are significantly smaller than those under static conditions[20,24,27]; and (3) the peripheral or near-track lesions within the on-track group are at risk of recurrent instability with a soft tissue repair.[4,17,18,28,29] To overcome these disadvantages of the radiologic track method, Bhatia et al[8] evaluated a Dynamic Arthroscopic Standardized Track (DAST) method and suggested that the highly reliable DAST could be incorporated into current decision-making algorithms to reduce variability associated with the static nature of radiologic measurements. In the present study, we propose a modification of currently used shoulder instability scores (ISI score and GTIMS) by substituting the image-based component with a dynamic track method (ie, DAST) (Table I).

This study aims to assess the reliability, reproducibility, and diagnostic validity of a new dynamic track score, the Dynamic Track Instability Management Score (DTIMS), and to compare these with current gold standard scores (ISI score and GTIMS) for predicting the necessity of a bony reconstruction procedure. The hypothesis was that in comparison with the currently used shoulder instability scores, DTIMS will have greater reliability and interobserver agreement for bone defect assessment and in decision making for soft tissue vs. bony surgery.

Table I Dynamic Track Instability Management Score (DTIMS)

DTIMS prognostic factors	Score
Age at surgery, yr	
≤20	2
>20	0
Type of sport	
Contact or forced overhead	1
Other	0
Level of competition in sport	
Competitive	2
Recreational or none	0
Shoulder hyperlaxity	
Confirmed anterior or inferior hyperlaxity	1
Normal laxity	0
Dynamic track classification of bone loss using DAST	
On-track (central and peripheral)	0
Off-track	4
Total score	10

DAST, Dynamic Arthroscopic Standardized Test; *GTIMS*, Glenoid Track Instability Management Score; *ISIS*, Instability Severity Index Score. Non–bone loss prognostic factors are the same in DTIMS, GTIMS, and ISI. Bone loss assessment is performed using the dynamic track method (DAST).

Methods

Between February 2023 and October 2024, consecutive patients who presented with recurrent dislocations and subluxations were prospectively evaluated clinically and radiologically to confirm traumatic anterior shoulder instability. In addition to standard shoulder radiographs, all patients underwent both a magnetic resonance imaging and a CT scan and were further assessed intraoperatively by arthroscopic evaluation. Institutional permissions and institutional review board approval was obtained, and patients who consented to be a part of the study were retrospectively evaluated. We included patients with 2 or more episodes of recurrent dislocations and subluxations, with clinical symptoms and signs of anterior instability (positive apprehension and other tests). We excluded (1) patients who did not consent to participate in the study, follow-up, or opted for conservative treatment; (2) patients with painful instability (acute episodes, evolving stiffness); (3) patients with a failed previous soft tissue surgery or bone grafting procedure; (4) presence of concomitant full-thickness rotator cuff tear; and (5) patients with additional posterior or multidirectional instability.

Clinical profile and evaluation

Clinical data documented for each patient included gender, age at the time of undergoing surgery, number of instability episodes and mechanisms, type and level of sport played, and use of the extremity in forced overhead activity. Instability was assessed with shoulder apprehension tests for anterior and posterior directions. Ranges of motion were documented. Hyperlaxity was evaluated independently by 2 fellowship-trained shoulder surgeons, in the anterior direction (external rotation greater than 85° with the arm at the side) and in an inferior direction (positive Gagey hyperabduction test).[15] In addition, generalized joint hypermobility was identified using the Beighton criteria (with scores ≥4 indicating hypermobility).[5]

Radiologic evaluation

All patients included in the study underwent magnetic resonance imaging to assess the soft tissue injury from anterior instability. Bone defects were evaluated using standard radiographs (true anteroposterior view, and anteroposterior view in external rotation) and using a CT scan. The images were reanalyzed using DICOM software by 2 fellowship-trained shoulder surgeons with significant previous experience in radiologic measurement techniques. The patient was not identified at the time of radiologic analysis, and all measurements were performed in a time frame that differed from the time of surgery to avoid bias in measurements. Glenoid bone loss (GBL) was assessed on CT scans (3-dimensionally reconstructed en face sagittal views) by the best-fit circle method.[1,3,6,7] GBL was calculated as the diameter of the defect divided by the diameter of the best-fit circle and was converted to a percentage. The Hill-Sachs interval (HSI) length was measured at the point of maximum distance from the medialmost aspect of the HSL to a line drawn in the plane of the greater tuberosity (medial aspect) on 3-dimensionally reconstructed humeral head.[12,13,19] The glenoid track (GT) was calculated using the mathematical formula described by Di Giacomo et al[12] ($GT = 0.83 \times [D - d]$), in which D is the diameter of the glenoid best-fit circle and d is the width of anterior GBL). Based on this, HSLs were classified as on-track or off-track lesions.

Arthroscopic evaluation

All patients included in the study underwent dry arthroscopic evaluation in the beach-chair position. A standardized method of evaluation, DAST, described earlier by Bhatia et al[8] was used to classify defects into on-track (central and peripheral) and off-track lesions. Two experienced surgeons performed the test during surgery, and the interpretation was documented independently. DAST was performed as follows: In the beach chair position (40° inclination), the arm was supported using a standard arm-positioning device. Before surgical preparation and draping, the arm-positioning device was attached in a way that would facilitate shoulder movement up to 120°-130° of abduction and elevation, as well as 20°-30° of horizontal extension beyond the coronal plane. Using a posterior portal arthroscopic approach, the HSL was visualized by abducting the arm in the scapular plane. The orientation of HSL was aligned with the anterior glenoid rim (AGR) by changing the abduction angle in the scapular plane until the longitudinal axis of the HSL was parallel to the AGR. If necessary, the shoulder was further abducted, horizontally extended, or externally rotated to achieve parallel HSL-AGR alignment as a starting point (Fig. 1). Thereafter, the humeral head was displaced anteriorly using firm thumb pressure (defined as maximum pressure possible without physically moving the patient from the beach chair positioner) applied to the posterior aspect of the humeral head; the end-point was achieved when no further anterior displacement was possible. With the humeral head in maximum anterior displacement, the HSL was visualized and defined as on-track (central or peripheral) or off-track based on the criteria described by Bhatia et al[8] (ie, on-track central = HSL between bare area and AGR, on-track peripheral = HSL displaced up to AGR without "engagement," and off-track = HSL translated beyond GR and "engaging") (Fig. 1).

Instability scores

Using clinical and radiologic data, anterior shoulder instability scores (ISI score and GTIMS) were calculated prior to surgery. During arthroscopic assessment, data obtained from DAST was used to calculate DTIMS. All scores were calculated independently by 2 observers. Based on the scores obtained, a theoretical decision to perform either a soft tissue procedure (score ≤3) or a bony (Latarjet) procedure (score ≥4) was documented.[12,21-23]

Statistical analysis

The data were entered into Microsoft Excel, and statistical analysis was carried out in Stata version 14.0 software (StataCorp LP, College Station, TX, USA). Cohen kappa statistic, along with its 95% confidence intervals, was calculated to evaluate the agreement between 2 observers for DTIMS as well as for GTIMS and ISI score, in diagnosing on-track vs. off-track lesions and determining the need for bone or soft tissue surgery. Kappa statistics were also employed to assess interobserver agreement in identifying HSLs and loss of glenoid contour on radiographs. Inter-rater reliability for total scores was assessed using the intraclass correlation coefficient (ICC). Cohen κ values were interpreted as follows: 0 = no agreement, 0.1-0.2 = slight agreement, 0.21-0.4 = fair agreement, 0.41-0.6 = moderate agreement, 0.6-0.8 = substantial agreement, 0.81-0.99 = near perfect agreement, 1 = perfect agreement. ICC values were interpreted as follows: <0.5 = poor, 0.5-0.75 = moderate, 0.75-0.9 = good, >0.9 = excellent. Furthermore, kappa statistics were applied to evaluate the agreement between different methods (eg, DTIMS vs. GTIMS, DTIMS vs. ISI score, and GTIMS vs. ISI score) for diagnosing on-track and off-track lesions, as well as in surgical decision making for bony or soft tissue procedures. Lastly, the diagnostic validity of DTIMS was assessed against both ISI score and GTIMS, and sensitivity, specificity, positive and negative

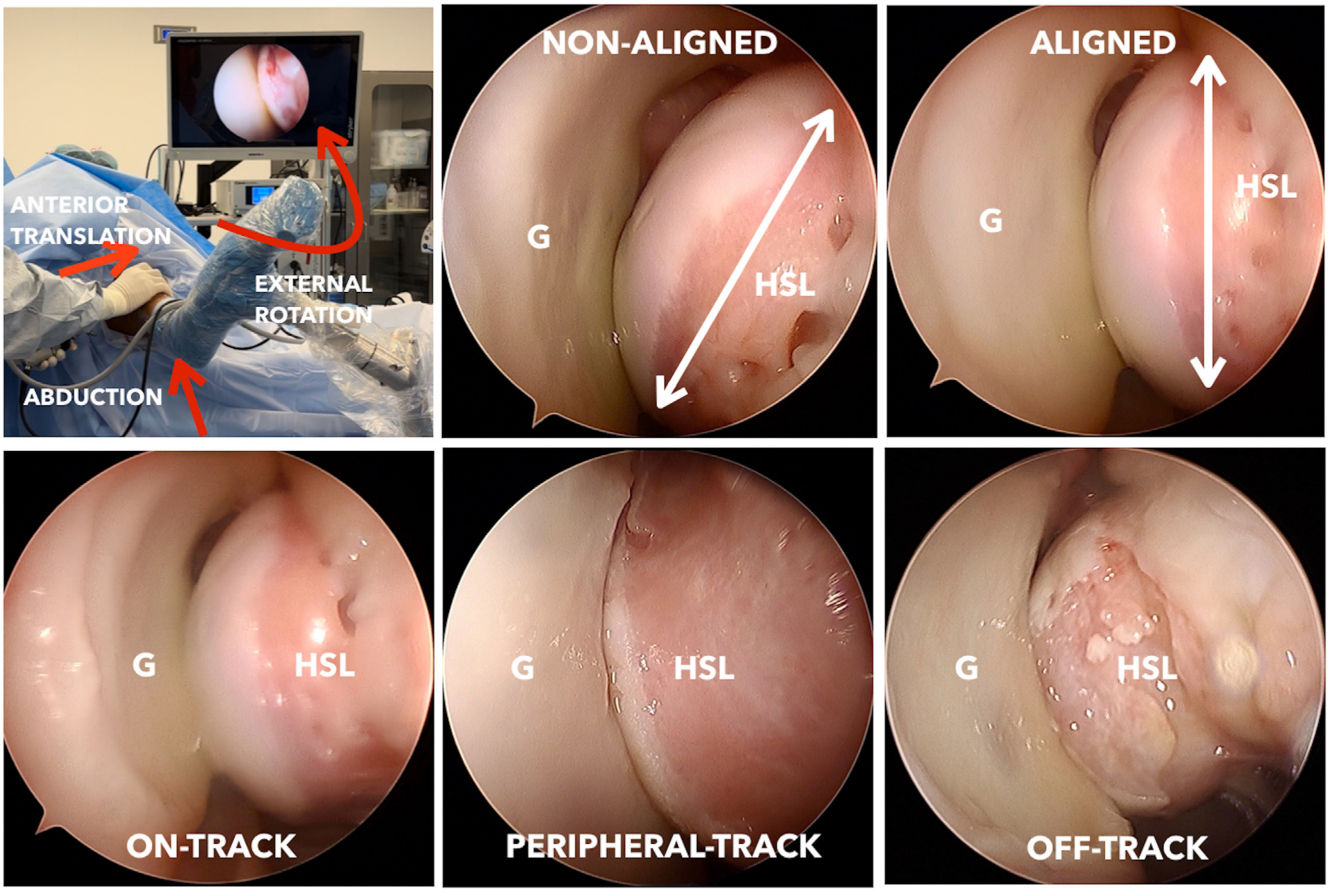

Figure 1 Intraoperative technique of performing the Dynamic Arthroscopic Standardized Test (DAST) is demonstrated. The shoulder is abducted in the scapular plane and externally rotated or extended to align the Hill-Sachs lesion (HSL) with the anterior glenoid (G), and the humeral head is translated anteriorly with manual pressure (*top* images). DAST is interpreted as follows: on-track central if the HSL is between the bare area and the anterior glenoid rim (*bottom left* image), on-track peripheral if the HSL is displaced up to the anterior glenoid rim without "engagement" (*bottom middle* image), and off-track if the HSL is translated beyond and engages the anterior glenoid rim (*bottom right* image).

predictive values, and diagnostic accuracy were reported. Similarly, the diagnostic validity of GTIMS was assessed against the ISI score, and the validity of the ISI score was evaluated against GTIMS. A P value of <.05 was considered statistically significant.

Results

Patient demographics

A total of 108 patients (14 female; 94 male) underwent radiologic imaging and arthroscopic evaluation for this study. There were 62 right and 46 left shoulders (64 dominant, 44 nondominant). Mean age at surgery was 26 years (range, 13-52 years). Thirty-eight patients were competitive athletes (contact or overhead sports), and 12 were recreational athletes. Thirty-four patients were found to be hyperlax, with a mean Beighton score of 6 (range 4-8). A bony procedure (Latarjet procedure) was performed in 68 patients, whereas arthroscopic soft tissue repair (anterior labral repair, labral repair and remplissage, and superior labrum anterior-posterior tear repair and/or biceps tenodesis) was performed in the remaining 40.

Interobserver reliability for bone loss assessment and decision making

The overall percentage agreement and κ values for identification of significant bone defects (radiographic [ISI score], CT scan [GTIMS], and dynamic [DTIMS]) are presented in Table II. Interobserver reliability for total score in each of the 3 methods was excellent (ICC >0.9). Interobserver reliability for identification of significant bone defects was near perfect ($\kappa = 0.94$, $P < .001$) with the dynamic method used in DTIMS and was substantial ($\kappa = 0.69$, $P < .001$) with the CT scan method used in GTIMS. However, reliability was lower (moderate, $\kappa = 0.47/0.58$, $P < .001$) with the radiographic method used in ISI score for significant Hill-Sachs and glenoid

Table II Interobserver agreement and reliability of radiographic (ISI), CT scan (GTIMS), and arthroscopic (DTIMS) methods for bone loss assessment and decision making in 3 instability scores

Bone loss assessment method	% agreement	*P* value for % agreement	95% CI, %	Interobserver reliability: κ	ICC for total score between 2 observers
Radiographic method (ISI)					0.92 (0.88-0.94)
Loss of inferior glenoid contour (true AP radiograph)	75	<.001*	65.8-82.8	0.47 (0.28-0.62) Moderate agreement	
Hill-Sachs lesion (AP radiograph in ER)	88	<.001*	80.3-93.4	0.58 (0.35-0.74) Moderate agreement	
Soft tissue vs. bony surgery	87	<.001*	79.2-92.7	0.74 (0.59-0.84) Substantial agreement	
CT scan method (GTIMS)					0.94 (0.91-0.96)
On- and off-track classification	90.7	<.001*	83.6-95.5	0.69 (0.49-0.83) Substantial agreement	
Soft tissue vs. bony surgery	92.3	<.001*	85.9-96.8	0.85 (0.71-0.92) Nearly perfect agreement	
Arthroscopic method using Dynamic Arthroscopic Standardised Test DAST (DTIMS)					0.97 (0.96-0.98)
On- and off-track classification	97.2	<.001*	92.1-99.4	0.94 (0.83-0.98) Nearly perfect agreement	
Soft tissue vs. bony surgery	95.4	<.001*	89.5-98.5	0.91 (0.79-0.96) Nearly perfect agreement	

ISI, Instability Severity Index; *CT*, computed tomography; *GTIMS*, Glenoid Track Instability Management Score; *DTIMS*, Dynamic Track Instability Management Score; *AP*, anteroposterior; *ER*, external rotation; *CI*, confidence interval; *ICC*, intraclass correlation coefficient.
* Statistically significant.

defects. Similarly, interobserver reliability for surgical decision making was near perfect for DTIMS (κ = 0.91, $P < .001$) and GTIMS (κ = 0.85, $P < .001$), and was lower (substantial, κ = 0.74, $P < .001$) for ISI score.

Interscore reliability for bone loss assessment and decision making

The overall percentage agreement and κ values (interscore comparison) for identification of on- and off-track bone defects (GTIMS and DTIMS) and for decision making between soft tissue and bony surgery for 3 scores (ie, ISI score, GTIMS, and DTIMS) are presented in Tables III-V. For both examiners, the DAST method identified approximately 36% (39 of 108) of lesions as off-track, and 66% (26 of 39) of these dynamic off-track lesions were identified as off-track by the CT method. Similarly, the CT scan method identified approximately 19% (21 of 108) as off-track, and 66% (13 of 21) of these radiologic off-track lesions were found to be off-track by the dynamic method. Overall, 57% (62 of 108) of patients were predicted to require a bony procedure during surgery when scored with DTIMS. In comparison, 43% (46 of 108) of patients were predicted to require a bony procedure when scored with GTIMS, and 45% (49 of 108) with ISI score. DTIMS showed only a fair agreement with GTIMS (κ = 0.24, $P < .003$) for diagnosis of on- and off-track defects, and moderate agreement with GTIMS (κ = 0.5, $P < .001$) and ISI score (κ = 0.4, $P < .001$) for surgical decision making (Tables II and III). Similarly, GTIMS showed moderate agreement (κ = 0.5, $P < .001$) for surgical decision making when compared with ISI score (Table IV).

Interscore diagnostic validity

DGTIMS was compared to GTIMS and ISI score, and GTIMS and ISI score were compared using either GTIMS or ISI score as the gold standard for prediction of a need for bony procedure for instability surgery. Diagnostic validity indices of the 4 combinations are shown in Table VI. DTIMS showed high sensitivity compared with GTIMS and ISI score, and ISI score showed high specificity compared with GTIMS.

Discussion

The most important finding of this study is that substituting the imaging-based component of ISI score and GTIMS with a dynamic track-based method in DTIMS showed near-perfect interobserver reliability for on- and off-track lesion identification and for predicting the need for soft tissue or bony procedure. Interscore agreement was moderate and comparable between all 3 scores. Our hypothesis

Table III Overall quantification and interscore reliability (percentage agreement and Cohen κ) of 2 scores (GTIMS and DTIMS) for identification of on-track vs. off-track lesion identification, and for decision making (soft tissue vs. bony surgery) for 2 observers

Scoring system (for on-track vs. on- and off-track identification)	GTIMS		Total	% agreement	Cohen κ	*P* value
	On	Off				
DTIMS						
Observer 1						
On	61	8	69			
Off	26	13	39	68.5 (58.9-77.1)	0.24 (0.02-0.40) (Fair agreement)	.003*
Total	87	21	108			
Observer 2						
On	63	7	70			
Off	26	12	38	69.4 (59.8-78.0)	0.24 (0.02-0.40) (Fair agreement)	.003*
Total	89	19	108			

Scoring system (for soft tissue vs. bony surgery)	GTIMS		Total	% agreement	Cohen κ	*P* value
	Soft tissue	Bony				
DTIMS						
Observer 1						
Soft tissue	41	5	46			
Bony	21	41	62	75.9 (66.8-83.6)	0.53 (0.34-0.66) (Moderate agreement)	<.001*
Total	62	46	108			
Observer 2						
Soft tissue	42	3	45			
Bony	22	41	63	76.9 (67.8-84.4)	0.55 (0.36-0.68) (Moderate agreement)	<.001*
Total	64	44	108			

GTIMS, Glenoid Track Instability Management Score; *DTIMS*, Dynamic Track Instability Management Score.
* Statistically significant.

was proved correct: when compared with the 2 other scores, DTIMS showed higher reliability for on- and off-track lesion identification, and equivalent (compared with GTIMS) or slightly better reliability (compared with ISI score) for predicting the need for soft tissue or bony procedure. Therefore, DTIMS is a reliable alternative to currently used scoring systems for decision making in anterior shoulder instability.

The ISI score was devised and validated by Balg and Boileau[2] and was the first ever instability score that could be used in an outpatient setting for surgical decision between soft tissue repair vs. a Latarjet procedure. The authors suggested that a bony procedure was indicated if the ISI score was ≥ 6 points. Further studies found that a lower threshold value of ≥ 4 points was more predictive of recurrent instability.[10,21,22,25] Jankaouskas et al[16] suggested that radiographic loss of the sclerotic anterior glenoid line is only moderately sensitive for glenoid bone loss. Dekker et al[10] found that using radiographs for identification of significant bone loss was inaccurate when compared with CT scan measurements and did not correlate with outcomes. Subsequently, Boileau and Balg[9] recommended that incorporating preoperative advanced imaging (CT scans or magnetic resonance imaging) to assess bone loss more accurately would complement the decision-making process when using ISI score. In the current study, we evaluated the reliability of bone loss assessment on radiographs and when using CT scans and a dynamic arthroscopic (ie, DAST) method; our results validate the findings of Dekker et al and show lower reliability of the radiographic method (moderate agreement) for humeral and glenoid bone loss when compared with the CT scan method for on- and off-track classification (substantial agreement). We found highest reliability (near perfect) with the dynamic track method (ie, DAST) for on- and off-track classification, and these results are similar to those described in a previous study by Bhatia et al.[8]

DiGiacomo et al[11,13] devised the CT scan–based glenoid track concept to represent combined humeral and glenoid

Table IV Overall quantification and inter-score reliability (percentage agreement and Cohen κ) of 2 scores (ISI and DTIMS) for decision making (soft tissue vs. bony surgery) for 2 observers

Scoring system (for soft tissue vs. bony surgery)				% agreement	Cohen κ	*P* value
	ISI		Total			
DTIMS	Soft tissue	Bony				
Observer 1						
Soft tissue	36	10	46			
Bony	23	39	62	69.4 (59.8-78.0)	0.40 (0.20-0.55) (Fair agreement)	<.001*
Total	59	49	108			
Observer 2						
Soft tissue	37	8	45			
Bony	26	37	63	68.5 (58.9-77.1)	0.39 (0.19-0.53) (Fair agreement)	<.001*
Total	63	45	108			

ISI, Instability Severity Index; *DTIMS*, Dynamic Track Instability Management Score.
* Statistically significant.

Table V Overall quantification and interscore reliability (percentage agreement and Cohen κ) of 2 scores (ISI and GTIMS) for decision making (soft tissue vs. bony surgery) for 2 observers

Scoring system (for soft tissue vs. bony surgery)				% agreement	Cohen κ	*P* value
	ISI		Total			
GTIMS	Soft tissue	Bony				
Observer 1						
Soft tissue	46	16	62			
Bony	13	33	46	73.2 (63.8-81.2)	0.46 (0.27-0.61) (Moderate agreement)	<.001*
Total	59	49	108			
Observer 2						
Soft tissue	51	13	64			
Bony	12	32	44	51.5 (67.8-84.4)	0.52 (0.34-0.67) (Moderate agreement)	<.001*
Total	63	45	108			

GTIMS, Glenoid Track Instability Management Score; *ISI*, Instability Severity Index.
* Statistically significant.

Table VI Diagnostic validity parameters of 3 instability scores for predicting bony surgery in anterior shoulder instability

Parameter for diagnostic validity	DTIMS vs. GTIMS (GTIMS = gold standard)		DTIMS vs. ISI (ISI = gold standard)		ISI vs. GTIMS (GTIMS = gold standard)		GTIMS vs. ISI (ISI = gold standard)	
	Observer 1	Observer 2	Observer 1	Observer 2	Observer 1	Observer 2	Observer 1	Observer 2
Sensitivity, %	89	93	80	82	67	71	72	73
Specificity, %	66	66	61	59	78	81	74	80
Positive predictive value, %	66	65	63	59	72	73	67	71
Negative predictive value, %	89	93	78	82	74	80	78	81
Diagnostic accuracy, %	76	77	69	66	73	77	73	77

DTIMS, Dynamic Track Instability Management Score; *GTIMS*, Glenoid Track Instability Management Score; *ISI*, Instability Severity Index.

bone loss as on- and off-track lesions. Thereafter, to overcome the inaccuracy of radiographic views in ISI score, they incorporated radiologic glenoid track into ISI score and suggested a modified ISI score named glenoid track instability management score (GTIMS). In a clinical study, they found that the ISI score showed a 2-fold increase in predicting a Latarjet in comparison to GTIMS. However, the overall rate of recurrent instability in the GTIMS group after arthroscopic Bankart repair was almost twice (8%) that of the ISI score cohort (4.5%).[12] Accuracy and reliability of the CT method of on- and off-track classification have been questioned recently in literature, and the importance of dynamic factors like joint laxity and capsulolabral tissue quality that influence humeral head translation have been highlighted.[8,14,20,24,26-28] In order to overcome the drawbacks of radiologic methods of bone loss assessment, a standardized dynamic method (DAST) of classification of bone defects into on- and off-track lesions was investigated and found to be highly reliable when compared with radiologic measurements.[8] In the current study, we modified the GTIMS and ISI score by substituting imaging-based bone defect classification with the DAST-based on- and off-track classification. DTIMS identified twice the number of off-track lesions (36%) in comparison with GTIMS (19%). Not all radiologic off-track lesions were identified as off-track on arthroscopy, and vice versa; this may be because some of these lesions might have been interpreted as peripheral-track morphology and some peripheral-track lesions as off-track by either method. The number of patients predicted to need a bony procedure was marginally higher with DTIMS (57%) when compared with GTIMS (43%) and ISI score (45%), and the reliability for predicting a bony procedure was near perfect with DTIMS (0.91) and GTIMS (0.85), and lower with ISI score (0.74). Notably, this study did not identify a difference in the number of predicted bony procedures between GTIMS and ISI score as described in a previous study.[12] Lastly, for predicting a bony procedure, DTIMS showed maximum sensitivity when compared with either GTIMS (89%, 93%) or ISI score (80%, 82%), whereas the ISI score demonstrated maximum specificity (78%, 81%) when compared with GTIMS.

DTIMS has several advantages over image-based scores: (1) Dynamic tracking is an independent method that is technically simple to perform in an intraoperative setting; however, although DAST is standardized, it is still a subjective method and preliminary training is necessary to be able to perform it correctly. (2) DTIMS does not need complex calculations, and radiation exposure from CT scanners can be avoided; however, preoperative imaging is crucial to understand the pathology prior to surgery. (3) DAST has near-perfect interobserver reliability and agreement, and this adds to the reliability of DTIMS. Finally, (4) DTIMS has high sensitivity, and this makes it an important tool for accurate decision making in high-risk patients (contact sport athletes, individuals with hyperlax shoulder joints). However, in clinical practice, preoperative imaging provides valuable information on the extent of bony defects and the on- and off-track nature of combined bone loss, and this is important for surgical planning and patient counseling. DTIMS should therefore be considered a useful adjunct to preoperative imaging-based decision making in anterior shoulder surgery, and not a substitute for it.

This study has limitations: First, a limited sample size of 108 patients was analyzed, and this number was based on previous studies with a similar sample size.[3,8] It is possible that the results may be different if a very large sample size is used. Second, dynamic tracking is subjective, and incorrect technique can result in altered results and false interpretation, thereby altering the score. Third, we were not able to assess intraobserver reliability in this study as the intraoperative DAST could be performed only once in each patient. Lastly, DTIMS is a new score and needs validation for clinical outcomes. Further studies are necessary to evaluate and compare each scoring system to determine the most accurate and reliable algorithm in the management of anterior shoulder instability.

Conclusion

Substituting the image-based identification of bone defects with a dynamic track method (ie, DAST) resulted in a highly reliable scoring system, the DTIMS, that can be used as an alternative to current gold standard scores (ISI score and GTIMS) for decision making in anterior shoulder instability.

Disclaimers:

Funding: No funding was disclosed by the authors.
Conflicts of interest: The authors, their immediate families, and any research foundations with which they are affiliated have not received any financial payments or other benefits from any commercial entity related to the subject of this article.

References

1. Arenas-Miquelez A, Dabirrahmani D, Sharma G, Graham PL, Appleyard R, Bokor DJ, et al. What is the most reliable method of measuring glenoid bone loss in anterior glenohumeral instability? A cadaveric study comparing different measurement techniques for glenoid bone loss. Am J Sports Med 2021;49:3628-37. https://doi.org/10.1177/03635465211041386
2. Balg F, Boileau P. The instability severity index score. A simple pre-operative score to select patients for arthroscopic or open shoulder stabilisation. J Bone Joint Surg Br 2007;89:1470-7. https://doi.org/10.1302/0301-620X.89B11.18962
3. Barchilon VS, Kotz E, Barchilon Ben-Av M, Glazer E, Nyska M. A simple method for quantitative evaluation of the missing area of the anterior glenoid in anterior instability of the glenohumeral joint.

Skeletal Radiol 2008;37(8):731-6. https://doi.org/10.1007/s00256-008-0506-8
4. Barrow AE, Charles SJ, Issa M, Rai AA, Hughes JD, Lesniak BP, et al. Distance to dislocation and recurrent shoulder dislocation after arthroscopic Bankart repair: rethinking the glenoid track concept. Am J Sports Med 2022;50:3875-80. https://doi.org/10.1177/03635465221128913
5. Beighton P, Solomon L, Soskolne CL. Articular mobility in an African population. Ann Rheum Dis 1973;32(5):413-8.
6. Bhatia DN, Kandhari V. How does anterior glenoid bone loss affect shoulder stability? A cadaveric analysis of glenoid concavity and bony shoulder stability ratio. J Shoulder Elbow Surg 2022;31:553-60. https://doi.org/10.1016/j.jse.2021.09.004
7. Bhatia DN, Kandhari V. Bone defect-induced alteration in glenoid articular surface geometry and restoration with coracoid transfer procedures: a cadaveric study. J Shoulder Elbow Surg 2019;28:2418-26. https://doi.org/10.1016/j.jse.2019.04.050
8. Bhatia DN, Malviya P. How does dynamic arthroscopic tracking compare with radiologic glenoid track for identification of on- and off-track lesions in anterior shoulder instability? J Shoulder Elbow Surg 2024;33(1):23-31. https://doi.org/10.1016/j.jse.2023.05.018
9. Boileau P, Balg F. Editorial commentary: should we condemn the shoulder instability severity index scoring system? Not at all!... can we improve its radiographic component? Yes, we can. Arthroscopy 2021;37(5):1392-6. https://doi.org/10.1016/j.arthro.2021.02.008
10. Dekker TJ, Peebles LA, Bernhardson AS, Golijanin P, Di Giacomo G, Hackett TR, et al. Limited predictive value of the instability severity index score: evaluation of 217 consecutive cases of recurrent anterior shoulder instability. Arthroscopy 2021;37(5):1381-91. https://doi.org/10.1016/j.arthro.2020.12.185
11. Di Giacomo G, Itoi E, Burkhart SS. Evolving concept of bipolar bone loss and the Hill-Sachs lesion: from "engaging/non-engaging" lesion to "on-track/off-track" lesion. Arthroscopy 2014;30(1):90-8. https://doi.org/10.1016/j.arthro.2013.10.004
12. Di Giacomo G, Peebles LA, Pugliese M, Dekker TJ, Golijanin P, Sanchez A, et al. Glenoid track instability management score: radiographic modification of the instability severity index score. Arthroscopy 2020;36:56-67. https://doi.org/10.1016/j.arthro.2019.07.020
13. Di Giacomo G, Piscitelli L, Pugliese M. The role of bone in glenohumeral stability. EFORT Open Rev 2018;3(12):632-40. https://doi.org/10.1302/2058-5241.3.180028
14. Funakoshi T, Hartzler RU, Stewien E, Burkhart SS. Hill-Sachs lesion classification by the glenoid track paradigm in shoulder instability: poor agreement between 3-dimensional computed tomographic and arthroscopic methods. Arthroscopy 2019;3:1743-9. https://doi.org/10.1016/j.arthro.2018.12.005
15. Gagey OJ, Gagey N. The hyperabduction test. J Bone Joint Surg Br 2001;83(1):69-74.
16. Jankauskas L, Rüdiger HA, Pfirrmann CW, Jost B, Gerber C. Loss of the sclerotic line of the glenoid on anteroposterior radiographs of the shoulder: a diagnostic sign for an osseous defect of the anterior glenoid rim. J Shoulder Elbow Surg 2010;19(1):151-6. https://doi.org/10.1016/j.jse.2009.04.013
17. Lee YJ, Kim C, Kim SJ, Yoon TH, Cho JY, Chun YM. Does an "off-track" Hill-Sachs lesion that is misclassified as "non-engaging" affect outcomes from Bankart repair alone compared with Bankart repair combined with remplissage? Arthroscopy 2021;37:450-6. https://doi.org/10.1016/j.arthro.2020.09.037
18. Li RT, Kane G, Drummond M, Golan E, Wilson K, Lesniak BP, et al. On-track lesions with a small distance to dislocation are associated with failure after arthroscopic anterior shoulder stabilization. J Bone Joint Surg Am 2021;103:961-7. https://doi.org/10.2106/JBJS.20.00917
19. Metzger PD, Barlow B, Leonardelli D, Peace W, Solomon DJ, Provencher MT. Clinical application of the "glenoid track" concept for defining humeral head engagement in anterior shoulder instability: a preliminary report. Orthop J Sports Med 2013;1(2):2325967113496213. https://doi.org/10.1177/2325967113496213
20. Omori Y, Yamamoto N, Koishi H, Futai K, Goto A, Sugamoto K, et al. Measurement of the glenoid track in vivo as investigated by 3-dimensional motion analysis using open MRI. Am J Sports Med 2014;42:1290-5. https://doi.org/10.1177/0363546514527406
21. Phadnis J, Arnold C, Elmorsy A, Flannery M. Utility of the instability severity index score in predicting failure after arthroscopic anterior stabilization of the shoulder. Am J Sports Med 2015;43(8):1983-8. https://doi.org/10.1177/0363546515587083
22. Rosenberg SI, Padanilam SJ, Pagni BA, Tjong VK, Sheth U. A lower Instability Severity Index score threshold may better predict recurrent anterior shoulder instability after arthroscopic Bankart repair: a systematic review. J ISAKOS 2021;6(5):295-301. https://doi.org/10.1136/jisakos-2020-000584
23. Rouleau DM, Hébert-Davies J, Djahangiri A, Godbout V, Pelet S, Balg F. Validation of the instability shoulder index score in a multicenter reliability study in 114 consecutive cases. Am J Sports Med 2013;41(2):278-82. https://doi.org/10.1177/0363546512470815
24. Schneider AK, Hoy GA, Ek ET, Rotstein AH, Tate J, Taylor DM, et al. Interobserver and intraobserver variability of glenoid track measurements. J Shoulder Elbow Surg 2017;26:573-9. https://doi.org/10.1016/j.jse.2016.09.058
25. Verweij LPE, van Spanning SH, Grillo A, Kerkhoffs GMMJ, Priester-Vink S, van Deurzen DFP, et al. Age, participation in competitive sports, bony lesions, ALPSA lesions, > 1 preoperative dislocations, surgical delay and ISIS score > 3 are risk factors for recurrence following arthroscopic Bankart repair: a systematic review and meta-analysis of 4584 shoulders. Knee Surg Sports Traumatol Arthrosc 2021;29(12):4004-14. https://doi.org/10.1007/s00167-021-06704-7
26. Vopat ML, Hermanns CA, Midtgaard KS, Baker J, Coda RG, Cheema SG, et al. Imaging modalities for the glenoid track in recurrent shoulder instability: a systematic review. Orthop J Sports Med 2021;9:23259671211006750. https://doi.org/10.1177/23259671211006750
27. Wu C, Wang Y, Wang C, Chen J, Xu J, Yu W, et al. Glenoid track width is smaller under dynamic conditions: an in vivo dual-fluoroscopy imaging study. Am J Sports Med 2022;50:3881-8. https://doi.org/10.1177/03635465221126650
28. Yamamoto N, Shinagawa K, Hatta T, Itoi E. Peripheral-track and central-track Hill-Sachs lesions: a new concept of assessing an on-track lesion. Am J Sports Med 2020;48:33-8. https://doi.org/10.1177/0363546519886319
29. Yang TC, Chen KH, Chiang ER, Chang MC, Ma HL. Using the "Hill-Sachs interval to glenoid track width ratio" for prediction of recurrent instability after arthroscopic Bankart repair. Orthop Traumatol Surg Res 2018;104:797-801. https://doi.org/10.1016/j.otsr.2018.02.013