JB&JS OPEN ACCESS

EVIDENCE-BASED SYSTEMATIC REVIEWS

Intrarater and Inter-rater Reliability of Tibial Plateau Fracture Classifications

Systematic Review and Meta-Analysis

Fardis Vosoughi, MD*, Iman Menbari Oskouie, MD, MPH*, Nazanin Rahimdoost, MD, Amir Kasaeian, PhD, and Arash Sherafat Vaziri, MD

Background: The interobserver and intraobserver reliability of various tibial plateau fracture (TPF) classifications has been examined in recent literature using radiography, computed tomography, and magnetic resonance imaging. The question remains as to which classification system provides the highest reliability. In this systematic review, we are going to evaluate the overall interobserver and intraobserver reliability of various TPF classifications in different imaging modalities.

Methods: We conducted a systematic review following Preferred Reporting Items for Systematic Reviews and Meta-Analyses guidelines. In February 2023, predefined terms were used for database search (Embase, PubMed, Scopus, Cochrane, and Web of Science). Meta-analysis of intrarater and inter-rater kappa coefficients was performed for each of the classifications in each modality.

Results: Thirty-four studies were included in this review. Schatzker's classification was more frequently used than others. It had a better intrarater kappa coefficient than the Hohl and Moore and Arbeitsgemeinschaft für Osteosynthesefragen/Orthopedic Trauma Association (AO/OTA) classifications in radiography ($\kappa = 0.72$, 95% confidence interval [CI] = 0.67-0.76, $p < 0.01$). The Schatzker and AO/OTA classifications had similar inter-rater reliability in the radiography modality ($\kappa = 0.53$, 95% CI = 0.51-0.54, $p < 0.01$; $\kappa = 0.53$, 95% CI = 0.5-0.55, $p < 0.01$; respectively). In 3-dimensional computed tomography, the Luo classification system showed the highest intrarater ($\kappa = 0.85$, 95% CI = 0.35-0.66) and inter-rater ($\kappa = 0.77$, 95% CI = 0.73-0.81) kappa coefficients.

Conclusion: Three-column classification proposed by Luo et al. was able to reach the highest degree and was the only classification with near-excellent inter-rater reliability.

Tibial plateau fractures (TPFs) pose a threat to the knee's perfect functioning. They can result from axial compressive pressures acting on the knee joint exclusively or in conjunction with varus or valgus stress[1]. 1% to 2% of all fractures and approximately 8% of fractures in elderly patients are TPFs[2].

Classification systems are used to (1) communicate in clinical practice, (2) provide preoperative planning recommendations, (3) make a prognosis of injury, and (4) research on aspects of the injury[3-9]. The OTA/AO[7,10], Schatzker[8], and Hohl[4] classification systems are the 3 most often used classification systems in the evaluation of TPF based on radiography.

*These authors contributed equally.

The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

The authors all agree for submission and publication of the manuscript.

The project was approved by the ethics committee of our university (Tehran University of medical sciences).

Disclosure: The **Disclosure of Potential Conflicts of Interest** forms are provided with the online version of the article (http://links.lww.com/JBJSOA/A665).

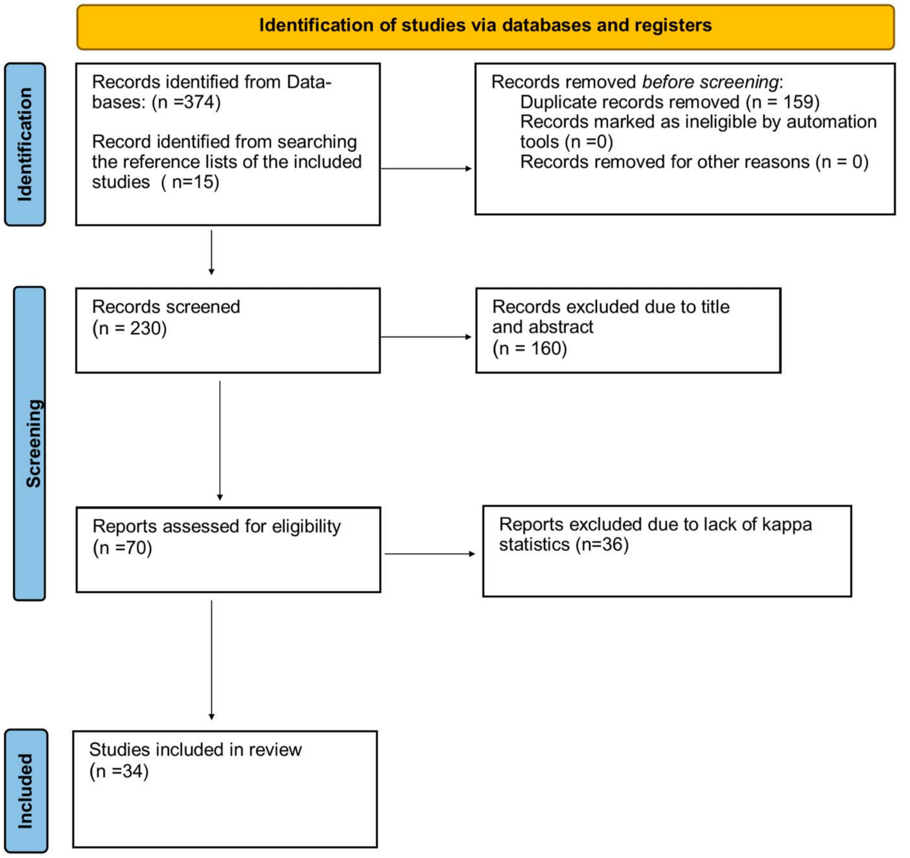

Fig. 1

PRISMA flow diagram for the systematic review detailing the database searches and number of abstracts screened and full texts reviewed. PRISMA = Preferred Reporting Items for Systematic Reviews and Meta-Analyses.

Currently, computed tomography (CT) scan is widely used because of its ability to accurately determine fracture lines and patterns, which provides surgeons with reliable information and leads to more accurate preoperative planning and fragment-specific fixation[11-17]. The "3-column classification" and "3-column fixation concept" were suggested by Luo et al. (2010) based on CT scans for TPF classifications[18]. The classification system proposed by Tunçez et al. for TPFs in 2022 used magnetic resonance imaging (MRI) and further assessment of soft-tissue injuries for TPF classifications[19].

The interobserver and intraobserver reliability of various TPF classifications describes the agreement among observers in classifying each TPF using common classification systems (e.g., Schatzker). This has been examined in recent literature using anterior-posterior and lateral radiographs[3,20,21], CT[22,23], and MRI[24]. To our best knowledge, no meta-analysis has been performed to estimate the overall kappa statistics of each classification/imaging modality. The question remains as to which classification system provides the most reliable method with the least interobserver/intraobserver variability.

In this systematic review, we are going to evaluate the overall interobserver and intraobserver reliability of various TPF classifications in different imaging modalities.

Methods

This report is a systematic review that is accorded with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines.

Protocol and Registration

Our article was registered through International Prospective Register of Systematic Reviews under the registration number of CRD42023461462.

Literature Search and Inclusion Criteria

All studies discussing TPFs and using a classification tool, either introducing a new classification system or modifying a preexisting one with inter-rater or intrarater kappa reliability of that classification, were included. Case reports, case series, letters to the editor, and any kind of reviews were excluded in the screening step.

Studies investigating TPF but not reporting fracture classification and kappa coefficients were excluded from the review.

We assessed the eligibility of study reports irrespective of their language, publication status, and publication date. Only articles published in peer-reviewed journals were included.

TABLE I Characteristics of the Contributing Studies*

Author	Year	Classification System	Imaging	Method	Average Intraobserver Kappa (k)				Average Interobserver Kappa (k)				
					XR	XR/2D-CT	XR/3D-CT	XR/3D Print	XR	XR/2D-CT	XR/3D-CT	XR/MRI	XR/3D Print
Alexander Brunner[22]	2010	OTA/AO Schatzker Hohl	XR XR/2D-CT	4 Observers 45 Fractures	AO/OTA: 0.619 Schatzker: 0.669 Hohl: 0.54	AO/OTA: 0.736 Schatzker: 0.695 Hohl: 0.779			AO/OTA: 0.429 Schatzker: 0.418 Hohl: 0.434	AO/OTA: 0.728 Schatzker: 0.755 Hohl: 0.771			
Peter Chan[42]	1997	Schatzker	XR XR/2D-CT	6 Observers 21 Fractures	Schatzker: 0.7	Schatzker: 0.8			Schatzker: 0.62	Schatzker: 0.61			
CP Charalambous[3]	2007	OTA/AO Schatzker	XR	6 Observers 50 Fractures	AO/OTA: 0.53 Schatzker: 0.57				AO/OTA: 0.43 Schatzker: 0.41				
Job N. Doornberg[15]	2011	OTA/AO Schatzker Hohl and Moore	XR/2D-CT XR/3D-CT	6 Observers 45 Fractures		AO/OTA: 0.723 Schatzker: 0.758 Hohl and Moore: 0.75	AO/OTA: 0.765 Schatzker: 0.746 Hohl and Moore: 0.814			AO/OTA: 0.536 Schatzker: 0.545 Hohl and Moore: 0.668	AO/OTA: 0.545 Schatzker: 0.596 Hohl and Moore: 0.75		
T.Gicquel[43]	2013	OTA/AO Schatzker Duparc	XR XR/2D-CT	6 Observers 50 Fractures	AO/OTA: 0.582 Schatzker: 0.626 Duparc: 0.663	AO/OTA: 0.694 Schatzker: 0.66 Duparc: 0.784			AO/OTA: 0.357 Schatzker: 0.404 Duparc: 0.365	AO/OTA: 0.479 Schatzker: 0.476 Duparc: 0.474			
Yan-Ling Hu[14]	2009	OTA/AO Schatzker	XR/2D-CT XR/3D-CT	4 Observers 21 Fractures		AO/OTA: 0.702 Schatzker: 0.756	AO/OTA: 0.887 Schatzker: 0.89			AO/OTA: 0.706 Schatzker: 0.736	AO/OTA: 0.828 Schatzker: 0.854		
Subramanyam Naidu Maripuri[20]	2008	OTA/AO Schatzker Hohl and Moore	XR	4 Observers 50 Fractures	AO/OTA: 0.8 Schatzker: 0.91 Hohl: 0.76				AO/OTA: 0.36 Schatzker: 0.47 Hohl: 0.14				
James Martin[44]	2000	OTA/AO	XR XR/2D-CT	5 Observers 56 Fractures					AO/OTA: 0.7	AO/OTA: 0.74			
Jos J. Mellema[45]	2016	Schatzker Luo	2D-CT	42 Observers 15 Fractures						Schatzker: 0.37 Luo: 0.31			
Jos J. Mellema[45]	2016	Schatzker Luo	2D-CT/3D-CT	39 Observers 15 Fractures							Schatzker: 0.29 Luo: 0.25		
Patange Subba Rao[46]	2014	Schatzker Luo	XR/2D-CT XR/2D-CT/3D-CT	5 Observers 56 Fractures		Schatzker: 0.624 Luo: 0.792	Schatzker: 0.68 Luo: 0.826			Schatzker: 0.536 Luo: 0.718	Schatzker: 0.552 Luo: 0.874		
Martijn A. J. te Stroet[47]	2011	Schatzker	XR XR/2D-CT	8 Observers 15 Fractures	Schatzker: 0.6	Schatzker: 0.57			Schatzker: 0.47	Schatzker: 0.46			
N. P. Walton[21]	2003	OTA/AO Schatzker	XR	3 Observers 30 Fractures	AO/OTA: 0.7 Schatzker: 0.68				AO/OTA: 0.41 Schatzker: 0.38				
David Wennergren[48]	2016	OTA/AO	XR	3 Observers 114Fractures	AO/OTA: 0.75				AO/OTA: 0.74				

continued

TABLE I (*continued*)

Author	Year	Classification System	Imaging	Method	Average Intraobserver Kappa (k)				Average Interobserver Kappa (k)				
					XR	XR/2D-CT	XR/3D-CT	XR/3D Print	XR	XR/2D-CT	XR/3D-CT	XR/MRI	XR/3D Print
Yi Zhu[49]	2013	OTA/AO Schatzker Luo	XR	4 Observers 50 Fractures					AO/OTA: 0.623 Schatzker: 0.567 Luo: 0.766				
Yi Zhu[50]	2012	Schatzker Luo	XR	4 Observers 278 Fractures	Schatzker: 0.758 Luo: 0.81				Schatzker: 0.567 Luo: 0.766				
H. Hoekstra[51]	2017	Schatzker Luo	XR/2D-CT	5 Observers 36 Fractures		Schatzker: 0.746 Luo: 0.782				Schatzker: 0.5 Luo: 0.675			
Rodrigo Pires e Albuquerque[52]	2009	Schatzker OTA/AO Hohl and Moore	XR	9 Observers 50 Fractures					Schatzker: 0.414 AO/OTA: 0.435 Hohl: 0.438				
Tom J. Crijns[53]	2020	Schatzker Luo OTA/AO	XR/3D-CT	23 Observers 20 Fractures							Schatzker: 0.28 Luo: 0.19 AO/OTA: 0.16		
Tom J. Crijns[53]	2020	Schatzker Luo OTA/AO	XR/3D-CT	26 Observers 20 Fractures							Schatzker: 0.35 Luo: 0.27 AO/OTA: 0.24		
Sageer Ahmad[23]	2022	Schatzker 4 quadrant	XR XR/3D-CT	5 Observers 35 Fractures	Schatzker: 0.7454		Schatzker: 0.7458 4 quadrant: 0.8892		Schatzker: 0.6048		Schatzker: 0.6141 Four quadrant: 0.8681		
Stephan Yacoubian[24]	2002	Schatzker	XR XR/2D-CT XR/MRI	3 Observers 52 Fractures					Schatzker: 0.68	Schatzker: 0.73		Schatzker: 0.85	
Jonatas Brito de Alencar Neto[54]	2020	Schatzker AO/ASIF Luo	XR XR/2D-CT	29 Observers 15 Fractures						Schatzker: 0.2345 AO/ASIF: 0.3653 Luo: 0.3161			
Panagiotis T.Masouros[55]	2022	Schatzker OTA/AO Luo	XR XR/3D-CT	12 Observers 25 Fractures					Schatzker: 0.361 AO/OTA: 0.204 Luo: 0.498		Schatzker: 0.364 AO/OTA: 0.231		
Mansoor Ali Khan[56]	2021	Schatzker KHan	XR	2 Observers 50 Fractures					Schatzker: 0.5 Khan: 0.26				

continued

TABLE I (continued)

Author	Year	Classification System	Imaging	Method	Average Intraobserver Kappa (k)				Average Interobserver Kappa (k)				
					XR	XR/2D-CT	XR/3D-CT	XR/3D Print	XR	XR/2D-CT	XR/3D-CT	XR/MRI	XR/3D Print
Peifeng Yao[57]	2022	Schatzker OTA/AO uTCC Ten segment	2D-CT 2D-CT/3D-CT	6 Observers 90 Fractures		Schatzker: 0.68 OTA/AO: 0.69 uTCC: 0.74 Ten segment: 0.8	Schatzker: 0.83 OTA/AO: 0.83 uTCC: 0.85 Ten segment: 0.91			Schatzker: 0.64 OTA/AO: 0.56 uTCC: 0.53 Ten segment: 0.6	Schatzker: 0.66 OTA/AO: 0.59 uTCC: 0.65 Ten segment: 0.73		
Anıl Taşkesen[58]	2017	Schatzker OTA/AO Hohl and Moore Duparc Luo	XR XR/2D-CT	4 Observers 60 Fractures	Schatzker: 0.585 OTA/AO: 0.485 Hohl and Moore: 0.57 Duparc: 0.51	Schatzker: 0.6625 OTA/AO: 0.59 Hohl and Moore: 0.66 Duparc: 0.6125 Luo: 0.725			Schatzker: 0.51 OTA/AO: 0.43 Hohl and Moore: 0.45 Duparc: 0.39	Schatzker: 0.61 OTA/AO: 0.54 Hohl and Moore: 0.51 Duparc: 0.52 Luo: 0.47			
B. Zhang[59]	2019	Schatzker TCC	XR/2D-CT	4 Observers 90 Fractures		Schatzker: 0.743 TCC: 0.724				Schatzker: 0.686 TCC: 0.739			
Alfredo Martínez-Rondanelli[11]	2017	Schatzker 4 column Duparc OTA/AO	XR/2D-CT	4 Observers 49 Fractures		Schatzker: 0.48 4 column: 0.61 Duparc: 0.37 AO: 0.34				Schatzker: 0.34 4 column: 0.34 Duparc: 0.23 AO: 0.11			
Tobias Dust[60]	2022	OTA/AO Ten segment Schatzker	XR/2D-CT XR/3D-CT	22 Observers 22 Fractures		OTA/AO: 0.42 Ten segment: 0.25 RevSchatzker: 0.38	OTA/AO: 0.42 Ten segment: 0.27 RevSchatzker: 0.41	OTA/AO: 0.47 Ten segment: 0.26 RevSchatzker: 0.47		OTA/AO: 0.32 Ten segment: 0.11 RevSchatzker: 0.24	OTA/AO: 0.32 Ten segment: 0.11 RevSchatzker: 0.28		OTA/AO: 0.34 Ten segment: 0.18 RevSchatzker: 0.31
Adeel Anwar[61]	2019	Schatzker Two column	XR/2D-CT	4 Observers 44 Fractures		Schatzker: 0.789 Two column: 0.955				Schatzker: 0.723 Two column: 0.939			
Angélica Millán-Billi[62]	2017	OTA/AO Schatzker Luo Deparc KHan	XR/3D-CT	4 Observers 112 Fractures			OTA/AO: 0.85 Schatzker: 0.87 Luo: 0.86 Deparc: 0.56 Khan: 0.43				OTA/AO: 0.62 Schatzker: 0.65 Luo: 0.73 Deparc: 0.37 Khan: 0.25		
Henrique Mansur[63]	2022	Schatzker Kfuri	XR/2D-CT	34 Observers 20 Fractures						Schatzker: 0.46 Kfuri: 0.3			

continued

TABLE I (*continued*)

Author	Year	Classification System	Imaging	Method	Average Intraobserver Kappa (k)				Average Interobserver Kappa (k)				
					XR	XR/2D-CT	XR/3D-CT	XR/3D Print	XR	XR/2D-CT	XR/3D-CT	XR/MRI	XR/3D Print
Jellina M. Huitema[64]	2022	Schatzker OTA/AO Luo	XR/3D-CT XR/3D print	25 Observers 40 Fractures							Schatzker: 0.514 OTA/AO: 0.359 Luo: 0.514		Schatzker: 0.539 OTA/AO: 0.372 Luo: 0.532
Robert Patzold[65]	2017	OTA/AO Schatzker	XR/3D-CT	4 Observers 81 Fractures							Schatzker: 0.785 OTA/AO: 0.72		
Marcello Teixeira Castiglia[66]	2018	Schatzker	XR XR/2D-CT XR/3D-CT	10 Observers 70 Fractures	Schatzker: 0.76	Schatzker: 0.75	Schatzker: 0.78		Schatzker: 0.58	Schatzker: 0.62	Schatzker: 0.64		

*2D-CT = 2-dimensional computed tomography, 3D-CT = 3-dimensional computed tomography, CT = computed tomography, MRI = magnetic resonance imaging, and TPF = tibial plateau fracture, XR = X-ray (plain radiography).

Information Sources

In September 2022, we searched Embase, PubMed, Scopus, Cochrane, and Web of Science for eligible studies, for all related English or non-English articles. We updated our database search in February 2023 using the same methods to ensure that the most recent studies were included in the review.

Search

Our search strategy is reported in line with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses literature search extension (PRISMA-S)[25]. The search strategy was reviewed by F.V. based on the Peer Review of Electronic Search Strategies[26] guideline. A detailed report of the search strategy is presented in Appendix A, http://links.lww.com/JBJSOA/A666.

Study Selection

We used EndNote software for citation management to import the citations extracted from literature searches and reference list checking[27]. EndNote's deduplication tool was used for identifying and removing similar data. After that, the remaining data were imported into Rayyan Qatar Computing Research Institute (QCRI)[28]. The first 50 records were given to reviewers (I.M.O. and N.R.) for independently reviewing the titles and abstracts of each article. Inter-rater reliability was calculated using Cohen's kappa to be 0.84, which is interpreted as almost perfect agreement.

In cases without agreement, another reviewer was employed (F.V.) for making the final decision. Later, for inclusion, the same 2 reviewers were employed to screen the full-text studies independently. If two independent raters agree on including a study, it will be included. However, if their opinions differ, a third reviewer will make the final decision[29].

Data Collection Process

A data extracting form was developed. In this form, 5 studies were randomly chosen to assess inter-rater reliability. This pilot test was performed by 2 reviewers. Evaluation of inter-rater reliability by Cohen's kappa at this stage was calculated to be 0.89, which shows perfect agreement. Any disagreements were discussed to be solved, for which the mentioned form was used separately by same reviewers. The new extracted data were compared. In cases with disagreement, details of additional discussions for resolving were entered. Final data were extracted in Microsoft Excel by the third reviewer (F.V.) and double-checked for probable mistakes[30].

Definitions for Data Extraction

The extracted data contained the name of the first author, year of publication, number of patients in the study, imaging modality used to introduce the classification, number of classes and subclasses (if present), average kappa value for intraobserver and/or interobserver reliability, and number of observers and fractures.

The reliability of inter-rater assessments, determined through independent evaluations of cases by 2 observers simultaneously, was used to demonstrate the level of agreement among raters. An evaluation of intrarater reliability, involving the same evaluator assessing cases at 2 different time points, was used to reveal the reproducibility of the results.

TABLE II Meta-Analysis of the Intrarater Kappa Coefficient*

Modality	Classification	Studies (n)	Observers (n)	Fractures (n)	Kappa Value	95% CI	p	Heterogeneity Tau^2	I^2	H	Egger's Test p-value
Radiography	Schatzker	11	60	2,631	0.7218	0.67-0.76	<0.01	0.1017	81.3%	2.32	0.748
	Hohl and Moore	3	12	620	0.6691	0.5-0.79	<0.01	0.1444	87.1%	2.79	0.068
	AO/OTA	7	30	1,400	0.648	0.58-0.71	<0.01	0.0609	73%	1.92	0.714
2D-CT	Schatzker	14	76	3,852	0.7202	0.69-0.74	<0.01	0.0473	67.4%	1.75	0.86
	Hohl and Moore	3	14	720	0.801	0.64-0.89	<0.01	0.3294	94%	4.08	0.376
	AO/OTA	8	56	2,324	0.6122	0.54-0.67	<0.01	0.13	83.4%	2.45	0.019*
	Luo	5	24	1,670	0.7573	0.73-0.78	<0.01	0.0061	32.7%	1.22	0.055
3D-CT	Schatzker	7	40	2,497	0.8129	0.77-0.85	<0.01	0.1032	82.9%	2.42	0.57
	AO/OTA	5	42	1856	0.7041	0.61-0.78	<0.01	0.2678	89.4%	3.07	<0.001*
	Luo	3	15	1,268	0.8491	0.83-0.86	<0.01	0	0	1	0.0888

*Significant publication bias. 2D-CT = 2-dimensional computed tomography, and 3D-CT = 3-dimensional computed tomography.

Risk of Bias

"Checklist for analytical cross-sectional studies" is a recommendation made by critical assessment tools for use in JBI Systematic Reviews. It is believed that this tool suits for evaluating bias[31,32] and is expressed as the most general tool used for assessing risk of bias and its application in systematic reviews of cross-sectional studies[33].

All included studies were assessed by the same 2 reviewers individually; the supporting information and justifications were recorded to assess risk of bias for each domain (low, unclear, high, or not applicable). For making the final decision, F.V. was employed.

Meta-Analysis

Meta-analyses were conducted using R version 4[34], function "SCSmeta"[35]. Hartung-Knapp adjustment for random-effects models was used for the meta-analysis. We used H statistics, Cochran's Q test, Higgin's and Thompson's I^2 statistics, and heterogeneity variance τ^2 statistics for measuring heterogeneity[36]. We also assessed heterogeneity by visual inspection of study data plots by using "Baujat plot". Forest plots were used to show the overall results of our meta-analysis.

Every study that was included had a number of observers, so there were multiple intrarater and inter-rater kappa coefficients recorded. As an example, a study with 4 observers found 6 inter-rater kappa coefficients and 4 intrarater ones. In our meta-analysis, we treated each of the mentioned values as a separate study.

Sensitivity Analysis

Sensitivity analysis was used to test the stability of our study. Sensitivity on modalities was analyzed. In addition, to evaluate the robustness of our results, at least 3 studies at low risk of bias in all modalities were involved.

TABLE III Meta-Analysis of the Inter-rater Kappa Coefficient*

Modality	Classification	Studies (n)	Pair of Observers (n)	Fractures (n)	Kappa Value	95% CI	p value	Heterogeneity Tau^2	I^2	H	Egger's Test p-value
Radiography	Schatzker	15	189	8,779	0.5262	0.51-0.54	<0.01	0.0065	10.4%	1.06	0.005*
	Hohl and Moore	4	39	1980	0.3973	0.35-0.44	<0.01	0.0032	13.5%	1.08	0.718
	AO/OTA	11	136	7,922	0.5268	0.50-0.55	<0.01	0.0148	45.1%	1.35	<0.001*
2D-CT	Schatzker	15	192	9,595	0.6061	0.59-0.62	<0.01	0.01	32%	1.21	0.005*
	Hohl and Moore	3	27	1,310	0.6687	0.61-0.72	<0.01	0.0326	59.8%	1.58	<0.001*
	AO/OTA	8	79	4,390	0.5745	0.54-0.61	<0.01	0.0329	61.4%	1.61	0.068
	Luo	5	46	3,134	0.627	0.59-0.66	<0.01	0.0263	64%	1.67	0.145
3D-CT	Schatzker	7	103	6,809	0.6573	0.64-0.67	<0.01	0.0043	30.2%	1.2	0.825
	AO/OTA	5	48	3,309	0.6218	0.60-0.64	<0.01	<0.001	15.1%	1.08	<0.001*
	Luo	3	36	2,862	0.7718	0.73-0.81	<0.01	0.0594	79.8%	2.23	<0.001*

*Significant publication bias. 2D-CT = 2-dimensional computed tomography, and 3D-CT = 3-dimensional computed tomography.

Study	Total	Weight	IV, Random, 95% CI
alexander brunner	45	1.8%	0.61 [0.39; 0.77]
alexander brunner	45	1.8%	0.49 [0.23; 0.69]
alexander brunner	45	1.8%	0.79 [0.64; 0.88]
alexander brunner	45	1.8%	0.78 [0.64; 0.88]
Peter Chan	21	1.4%	0.70 [0.38; 0.87]
Peter Chan	21	1.4%	0.70 [0.38; 0.87]
Peter Chan	21	1.4%	0.70 [0.38; 0.87]
Peter Chan	21	1.4%	0.70 [0.38; 0.87]
Peter Chan	21	1.4%	0.70 [0.38; 0.87]
Peter Chan	21	1.4%	0.70 [0.38; 0.87]
CP Charalambous	50	1.8%	0.53 [0.30; 0.70]
CP Charalambous	50	1.8%	0.63 [0.43; 0.77]
CP Charalambous	50	1.8%	0.70 [0.52; 0.82]
CP Charalambous	50	1.8%	0.36 [0.09; 0.58]
CP Charalambous	50	1.8%	0.51 [0.27; 0.69]
CP Charalambous	50	1.8%	0.66 [0.47; 0.79]
T.Gicquel	50	1.8%	0.63 [0.42; 0.77]
T.Gicquel	50	1.8%	0.63 [0.42; 0.77]
T.Gicquel	50	1.8%	0.63 [0.42; 0.77]
T.Gicquel	50	1.8%	0.63 [0.42; 0.77]
T.Gicquel	50	1.8%	0.63 [0.42; 0.77]
T.Gicquel	50	1.8%	0.63 [0.42; 0.77]
Subramanyam Naidu Maripuri	50	1.8%	0.76 [0.61; 0.86]
Subramanyam Naidu Maripuri	50	1.8%	0.92 [0.86; 0.95]
Subramanyam Naidu Maripuri	50	1.8%	0.99 [0.98; 0.99]
Subramanyam Naidu Maripuri	50	1.8%	0.97 [0.95; 0.98]
Martijn A. J. te Stroet	15	1.2%	0.60 [0.13; 0.85]
Martijn A. J. te Stroet	15	1.2%	0.60 [0.13; 0.85]
Martijn A. J. te Stroet	15	1.2%	0.60 [0.13; 0.85]
Martijn A. J. te Stroet	15	1.2%	0.60 [0.13; 0.85]
Martijn A. J. te Stroet	15	1.2%	0.60 [0.13; 0.85]
Martijn A. J. te Stroet	15	1.2%	0.60 [0.13; 0.85]
Martijn A. J. te Stroet	15	1.2%	0.60 [0.13; 0.85]
Martijn A. J. te Stroet	15	1.2%	0.60 [0.13; 0.85]
N. P. Walton	30	1.6%	0.68 [0.42; 0.84]
N. P. Walton	30	1.6%	0.68 [0.42; 0.84]
N. P. Walton	30	1.6%	0.68 [0.42; 0.84]
Yi Zhu	50	1.8%	0.85 [0.76; 0.91]
Yi Zhu	50	1.8%	0.75 [0.60; 0.85]
Yi Zhu	50	1.8%	0.74 [0.58; 0.84]
Yi Zhu	50	1.8%	0.69 [0.51; 0.81]
Sageer Ahmad	35	1.7%	0.76 [0.57; 0.87]
Sageer Ahmad	35	1.7%	0.70 [0.48; 0.84]
Sageer Ahmad	35	1.7%	0.74 [0.53; 0.86]
Sageer Ahmad	35	1.7%	0.70 [0.48; 0.84]
Sageer Ahmad	35	1.7%	0.84 [0.70; 0.91]
Anıl Taşkesen	60	1.8%	0.56 [0.36; 0.71]
Anıl Taşkesen	60	1.8%	0.60 [0.41; 0.74]
Anıl Taşkesen	60	1.8%	0.54 [0.33; 0.70]
Anıl Taşkesen	60	1.8%	0.64 [0.46; 0.77]
Marcello Teixeira Castiglia	70	1.9%	0.74 [0.61; 0.83]
Marcello Teixeira Castiglia	70	1.9%	0.85 [0.77; 0.90]
Marcello Teixeira Castiglia	70	1.9%	0.75 [0.63; 0.84]
Marcello Teixeira Castiglia	70	1.9%	0.78 [0.67; 0.86]
Marcello Teixeira Castiglia	70	1.9%	0.80 [0.70; 0.87]
Marcello Teixeira Castiglia	70	1.9%	0.72 [0.58; 0.82]
Marcello Teixeira Castiglia	70	1.9%	0.71 [0.57; 0.81]
Marcello Teixeira Castiglia	70	1.9%	0.81 [0.71; 0.88]
Marcello Teixeira Castiglia	70	1.9%	0.73 [0.60; 0.82]
Marcello Teixeira Castiglia	70	1.9%	0.75 [0.63; 0.84]
Total (95% CI)	**2631**	**100.0%**	**0.72 [0.68; 0.76]**

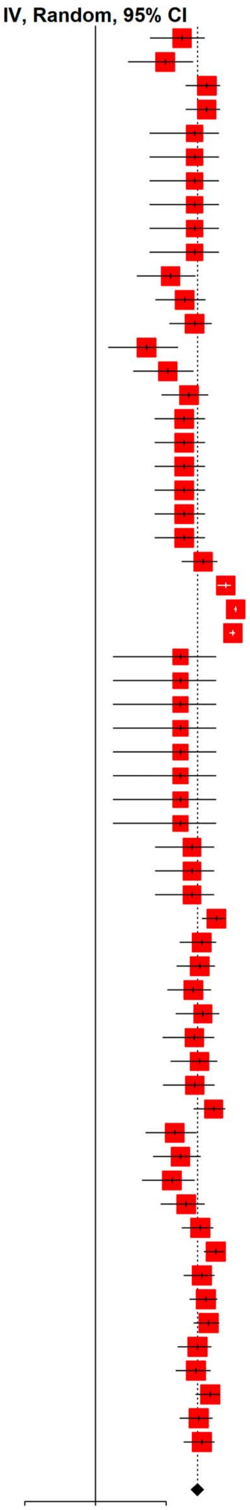

Fig. 2

Forest plot of intrarater kappa coefficients of the Schatzker classification in the radiography modality.

Results

Reporting Bias Assessment

There is a lack of systematic reviews regarding the impact of publication bias for evaluating whether the publication bias may interfere with the mean reliability coefficients acquired in the different meta-analyses. Egger tests and contour-enhanced funnel plots were conducted using the trim-and-fill method.

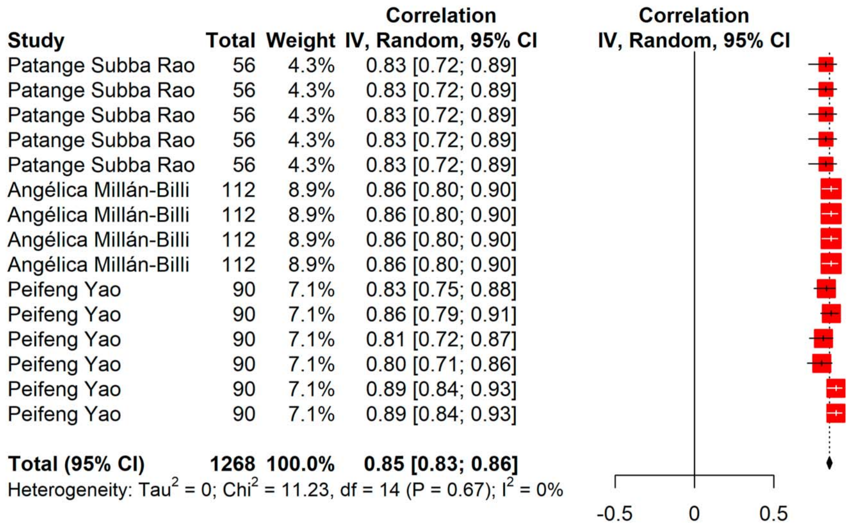

Fig. 3-A

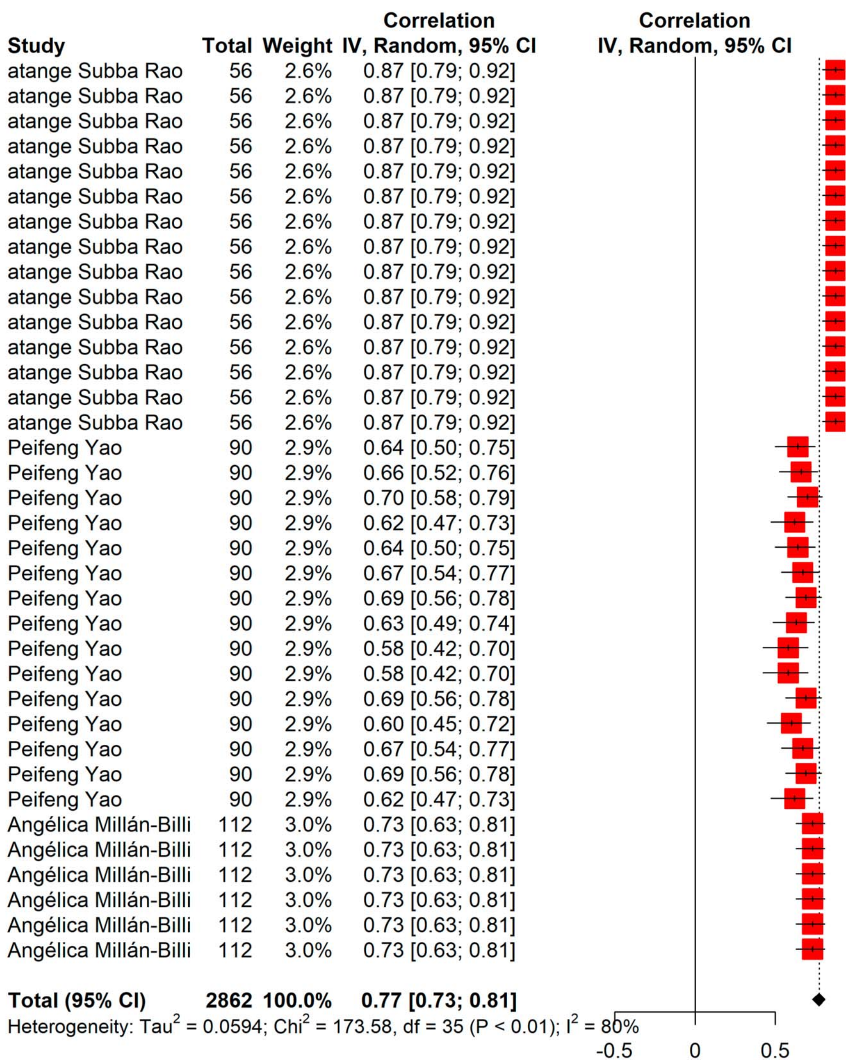

Fig. 3-B

Fig. 3 Forest plot of (**Fig. 3-A**) intrarater and (**Fig. 3-B**) inter-rater kappa coefficients of the Luo classification in the 3D-CT modality. 3D-CT = 3-dimensional computed tomography.

TABLE IV Trim-and-fill Test Results of the Intrarater Kappa Coefficient*

Modality	Classification	Total Observers (n)	Studies (Observers) Added (n)	Effect Size	95% CI	p	Heterogeneity			Cochrane p value
							Tau^2	I^2	H	
Radiography	Schatzker	87	27	0.76	0.73-0.79	<0.01	0.0926	80.8%	2.28	530.45
	Hohl and Moore	30	0	0.648	0.58-0.71	<0.01	0.0609	73%	1.92	107.37
	AO/OTA	12	0	0.6691	0.50-0.79	<0.01	0.1444	87.1%	2.79	85.34
2D-CT	Schatzker	76	11	0.7404	0.71-0.77	<0.01	0.0544	69.8%	1.82	284.47
	Hohl and Moore	14	0	0.8010	0.64-0.89	<0.01	0.3294	94.0%	4.08	216.31
	AO/OTA	76	20	0.6989	0.64-0.75	<0.01	0.1541	83.7%	2.48	461.21
	Luo	29	5	0.7365	0.70-0.77	<0.01	0.0168	51.4%	1.43	57.59
3D-CT	Schatzker	47	7	0.8363	0.80-0.87	<0.01	0.117	84.3%	2.52	293.2
	AO/OTA	55	13	0.8013	0.73-0.86	<0.01	0.3695	89.8%	3.14	531.21
	Luo	18	3	0.8568	0.84-0.87	<0.01	0.0006	4.6%	1.02	17.82

*2D-CT = 2-dimensional computed tomography, and 3D-CT = 3-dimensional computed tomography.

Study Selection

A total of 374 records were identified using the search strategy. An additional 15 records were identified by searching the reference lists of the included studies. After duplicates were removed, 230 records underwent Level 1 screening (screening by title/abstract) by Rayyan; 70 were subsequently assessed in Level 2 screening (screening by full text). Fifty-five were deemed relevant for data extraction, of which 34 (1-34) were included in our meta-analysis (Fig. 1).

Study Characteristics

A total of 1,892 fractures were studied by 384 observers across all 34 studies. A summary of the characteristics of the contributing studies is provided in Table I. All 34 studies included the Schatzker classification, 22 studies contained the AO/OTA classification, 5 studies included the Hohl and Moore classification, and 13 studies included the Luo classification.

The mean sample size of fractures in each study was 52.55 (interquartile range 21-56). The smallest sample size was 15 fractures, and the largest sample size was 278 fractures.

Primary Outcome

Separate meta-analyses were conducted as a function of type of reliability (intrarater and inter-rater) and type of modality (radiography, 2-dimensional computed tomography [2D-CT], and 3-dimensional computed tomography [3D-CT]). Thus, a total of 10 meta-analyses for intrarater and 10 meta-analyses for inter-rater reliability were conducted.

TABLE V Trim-and-fill Test Results of the Inter-rater Kappa Coefficient*

Modality	Classification	Total Pair of Observers (n)	Studies (Pair of Observers) Added (n)	Effect Size	95% CI	p	Heterogeneity			Cochrane p value
							Tau^2	I^2	H	
Radiography	Schatzker	249	60	0.5827	0.56-0.60	<0.01	0.019	43.4%	1.33	438.31
	Hohl and Moore	39	0	0.3973	0.35-0.44	<0.01	0.0562	13.5%	1.08	43.95
	AO/OTA	173	37	0.5788	0.56-0.60	<0.01	0.0299	61%	1.60	441.18
2D-CT	Schatzker	227	35	0.6224	0.60-0.64	<0.01	0.0133	34.8%	1.24	346.8
	Hohl and Moore	32	5	0.6240	0.55-0.69	<0.01	0.0726	75.9%	2.04	128.78
	AO/OTA	79	0	0.5745	0.53-0.61	<0.01	0.0329	61.4%	1.61	202.18
	Luo	59	13	0.5660	0.52-0.61	<0.01	0.0543	78.8%	2.17	272.95
3D-CT	Schatzker	117	14	0.6731	0.65-0.69	<0.01	0.009	40.5%	1.3	195.1
	AO/OTA	53	5	0.5971	0.54-0.65	<0.01	<0.001	57.1%	1.53	58.22
	Luo	43	7	0.7366	0.69-0.78	<0.01	0.0853	83.4%	2.45	252.51

*2D-CT = 2-dimensional computed tomography, and 3D-CT = 3-dimensional computed tomography.

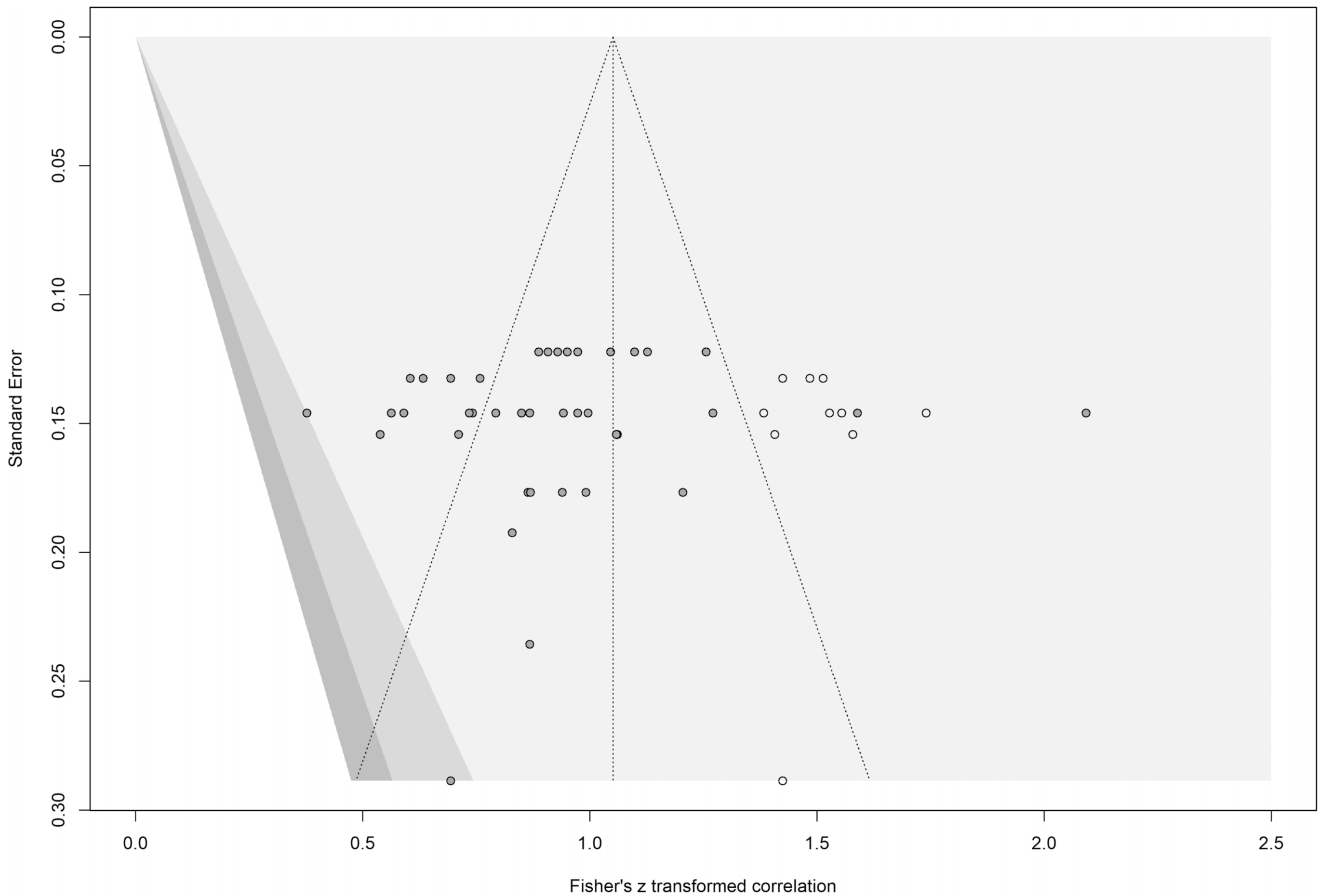

Fig. 4-A

Tables II and III present the results of each of the 20 meta-analyses conducted. Appendix B represents the forest plots of these analyses.

Intrarater Reliability

Intrarater kappa coefficients of 3 classifications (Schatzker, Hohl and Moore, and AO/OTA) with the radiography modality, 4 classifications (Schatzker, Hohl and Moore, AO/OTA, Luo) by using 2D-CT, and 4 classifications using 3D-CT (Schatzker, AO/OTA, Luo, 10) were calculated by meta-analysis in R. The Schatzker classification was more frequent than other classifications in all modalities. Eleven studies with the plain radiography modality used Schatzker classification for classifying their plateau fractures and it had the greatest intra-rater kappa coefficient in comparison with the Hohl and Moore and AO/OTA classifications (Schatzker effect size [ES] = 0.72, 95% CI = 0.67-0.76, $p < 0.01$, $I^2 = 81.3\%$; Hohl and Moore ES = 0.67, 95% CI = 0.5-0.79, $p < 0.01$, $I^2 = 87.1\%$; AO/OTA ES = 0.65, 95% CI = 0.58-0.71, $p < 0.01$, $I^2 = 73\%$) (Fig. 2).

By using 2D-CT as an imaging modality, Hohl and Moore had the greatest intrarater kappa coefficient (ES = 0.801, 95% CI = 0.64-0.89, $p < 0.01$, $I^2 = 94\%$). After that, the Luo classification also had a good intrarater kappa coefficient (ES = 0.7573, 95% CI = 0.73-0.78, $p < 0.01$, $I^2 = 32.7\%$). So we can deduce that Hohl and Moore and Luo classifications are more reliable than Schatzker based on 2D-CT (ES = 0.72, 95% CI = 0.69-0.74, $p < 0.01$, $I^2 = 67.4\%$).

In 3D-CT, Luo had the greatest intrarater kappa coefficient (ES = 0.85, 95% CI = 0.35-0.66, $p < 0.01$, $I^2 = 0\%$). By using the 3D-CT modality, an increase was seen in intrarater kappa coefficient for Luo, Schatzker (ES = 0.81, 95% CI = 0.77-0.85, $p < 0.01$, $I^2 = 82.9\%$), and AO/OTA (ES = 0.70, 95% CI = 0.61-0.78, $p < 0.01$, $I^2 = 89.4\%$), compared with other modalities (Fig. 3).

Appendix C represents Baujat plots that visualize overall heterogeneity contribution across influence on pooled result.

Inter-rater Reliability

Inter-rater reliability of 3 classifications (Schatzker, Hohl and Moore, AO/OTA) in the radiography modality, 4 classifications (Schatkzer, Hohl and Moore, AO/OTA, and Luo) in the 2D-CT modality, and 3 classifications (Schatzker, AO/OTA, Luo) in the 3D-CT modality were calculated by meta-analysis in R.

Schatzker and AO/OTA had similar inter-rater reliability in the radiography modality (ES = 0.53, 95% CI = 0.51-0.54, $p < 0.01$, $I^2 = 10.4\%$; ES = 0.53, 95% CI = 0.5-0.55, $p < 0.01$,

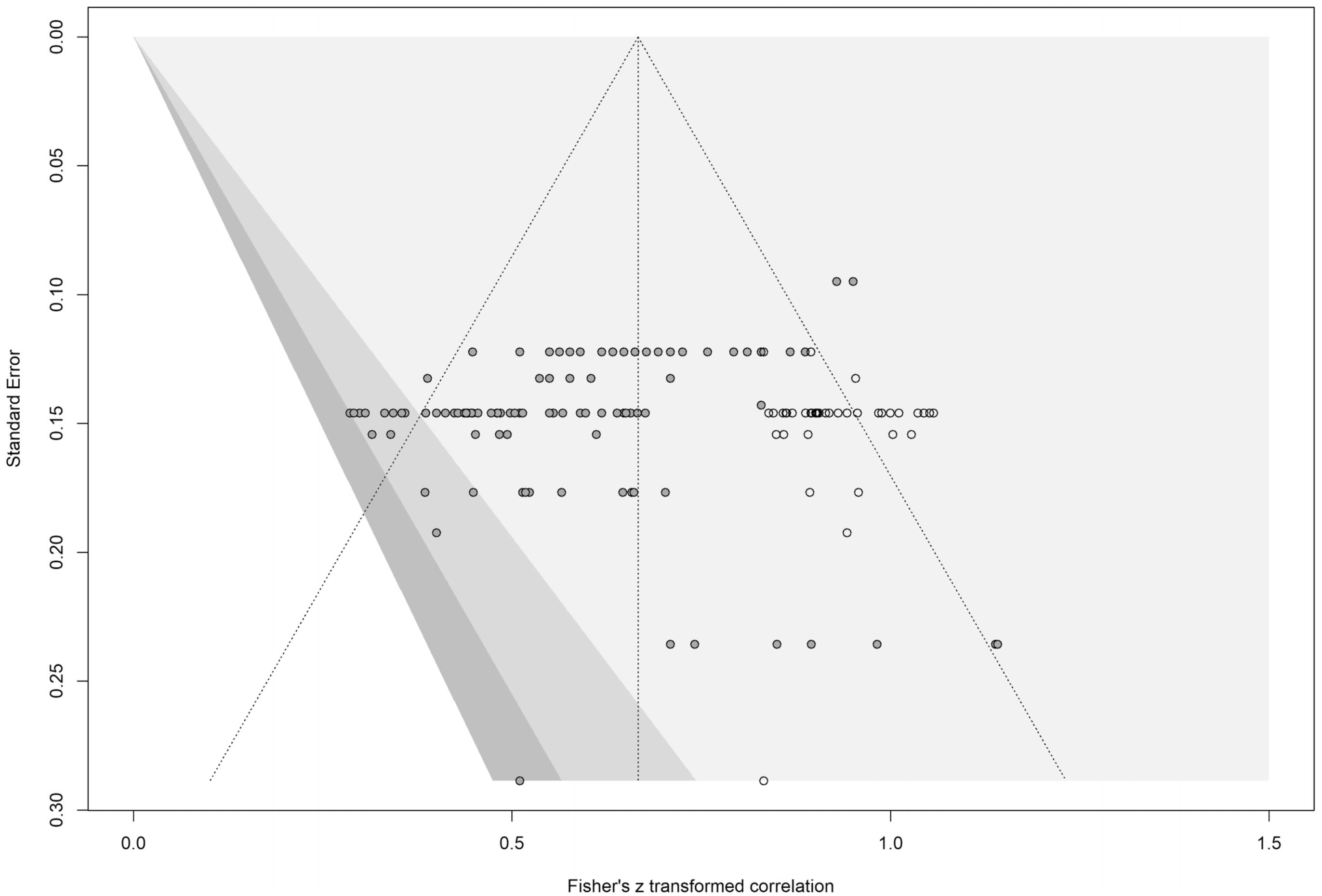

Fig. 4-B

Fig. 4 Funnel plot of (**Fig. 4-A**) intra-rater and (**Fig. 4-B**) inter-rater kappa coefficients of the Schatzker classification in the radiography modality.

I^2 = 45.1%; respectively), and they were more reliable than Hohl and Moore (ES = 0.40, 95% CI = 0.35-0.44, $p < 0.01$, I^2 = 13.5%).

In 2D-CT, the Hohl and Moore classification had the greatest inter-rater kappa coefficient (ES = 0.67, 95% CI = 0.61-0.72, $p < 0.01$, I^2 = 61.4%). We can see an increase in inter-rater kappa coefficients of all classifications by using 2D-CT compared with radiography.

In 3D-CT, the Luo classification had the greatest inter-rater kappa coefficient (ES = 0.77, 95% CI = 0.73-0.81, $p < 0.01$, I^2 = 61.4%) (Fig. 3). Again, we can see an increase of inter-rater kappa coefficient in Schatzker and AO/OTA in 3D-CT compared with 2D-CT.

Appendix C represents Baujat plots that visualize overall heterogeneity contribution across influence on pooled result.

Risk of Bias

Appendix D shows the risk of bias and concerns regarding applicability for each domain across the included studies. There were 12 studies with low risk of bias and 21 studies with moderate risk of bias. No study with high risk of bias was detected.

Publication Bias

After conducting an Egger's test, the result showed that there was publication bias in some meta-analyses of our study ($p < 0.05$); Tables IV and V present the results of the Egger test to each of the 20 meta-analyses. The absence of statistical significance for the intercept in all is a reason to discard publication bias. In addition, contour-enhanced funnel plots were constructed, and the trim-and-fill method for imputing missing values developed by Duval and Tweedie was applied. Tables IV and V present the results of the trim-and-fill method. Appendix E represents contour-enhanced funnel plots for each of the 20 meta-analyses. If the trim-and-fill method did not impute values to give symmetry to the funnel plot, it means that publication bias can be discarded as a threat to the meta-analytic results. The trim-and-fill test showed that there was article filling, indicating that there was publication bias or small sample bias in our included studies (Figs. 4 and 5)

Sensitivity Analysis

For evaluating sensitivity on reliability of various classifications in different modalities, studies with moderate or high risk of bias were omitted and meta-analyses were

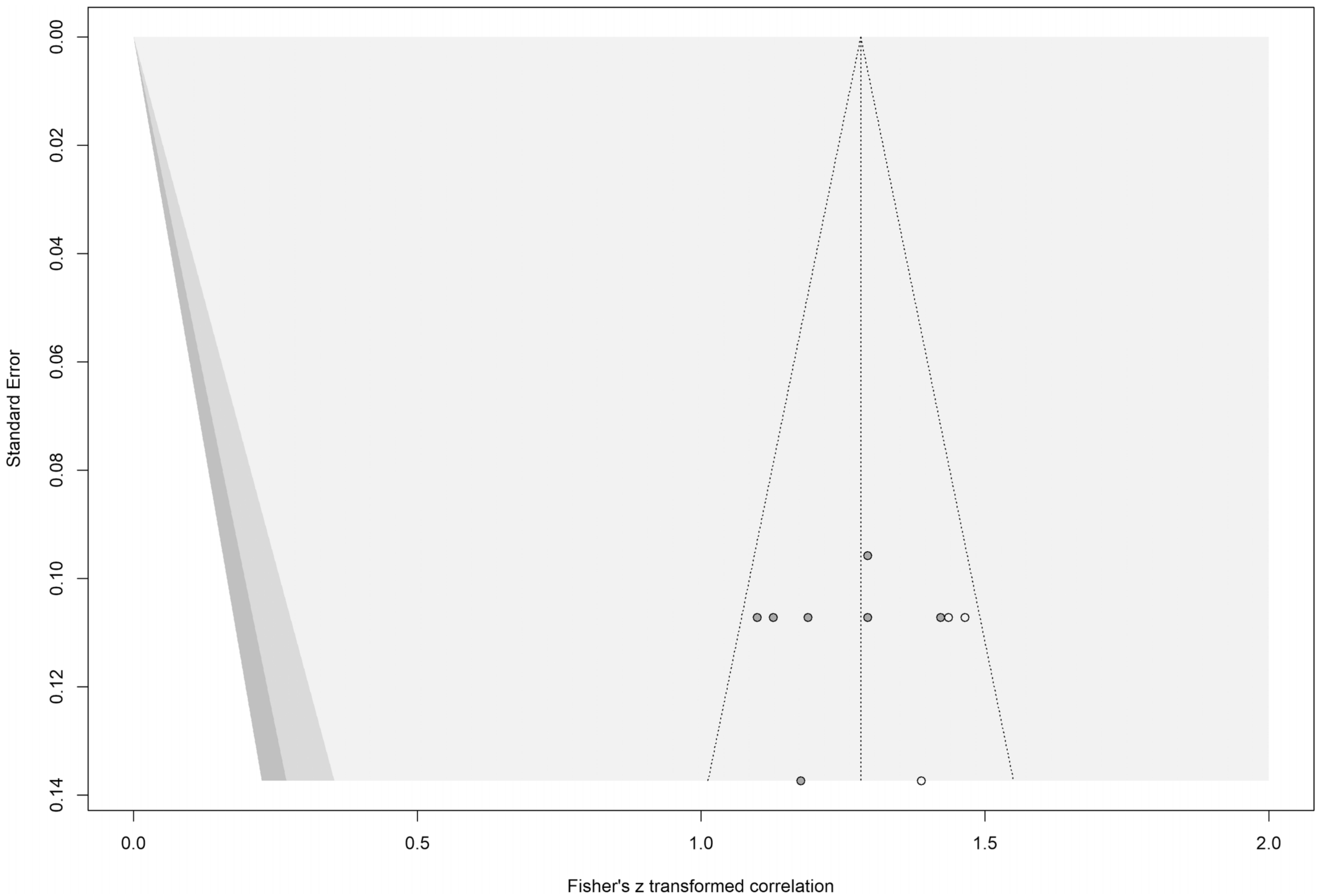

Fig. 5-A

conducted for at least 3 studies in each modality and classification.

In intrarater kappa coefficients (Appendix F, Supplementary Table 1), there was not any significant change for radiography. In 2D-CT, the kappa coefficient of AO/OTA increased from 0.61 to 0.73, but other classifications did not change. In 3D-CT again, improvement of the AO/OTA kappa coefficient was noted. There was not enough study for sensitivity analysis of the Luo classification in the 3D-CT modality.

In inter-rater kappa coefficients (Appendix F, Supplementary Table 2), like intrarater, there was no noticeable change in radiography imaging. In 2D-CT, sensitivity analysis was not possible for the Hohl and Moore classification, but in Schatzker and AO/OTA, a small increase in kappa coefficients was observed. In addition, the Luo classification had a small decrement. In 3D-CT, only the Schatzker classification was compatible for sensitivity analysis, and it did not change significantly.

Through the sensitivity analysis, the results showed that after excluding the moderate-bias studies, the results of the recombined data were almost unchanged, thus indicating that our study was stable and reliable.

Discussion

The aim of this study was to determine the most reliable classification for TPF. To our best knowledge, this is the first study to provide a comprehensive systematic review of TPF classification systems along with measuring their reliability and conducting meta-analysis for each classification in various modalities[37]. Millar et al.[38] conducted a systematic review of TPF and described the kappa values for different classifications, but they did not perform meta-analyses. In addition, we discovered several recently published articles to add to their review.

This review aimed to calculate the intrarater and inter-rater kappa coefficients of each classification in different imaging modalities and compare them together to find the most reliable classification in each modality. Landis and Koch criteria were used to interpret the kappa values (Table VI)[39].

The literature search identified 34 studies relating to TPF that reported inter-rater and intrarater reliability. Most of these classifications have not been used extensively in the literature, but they might be of value in clinical practice (Table I). Earlier systems presented basic explanations of fracture patterns, whereas more recent ones have tried to characterize

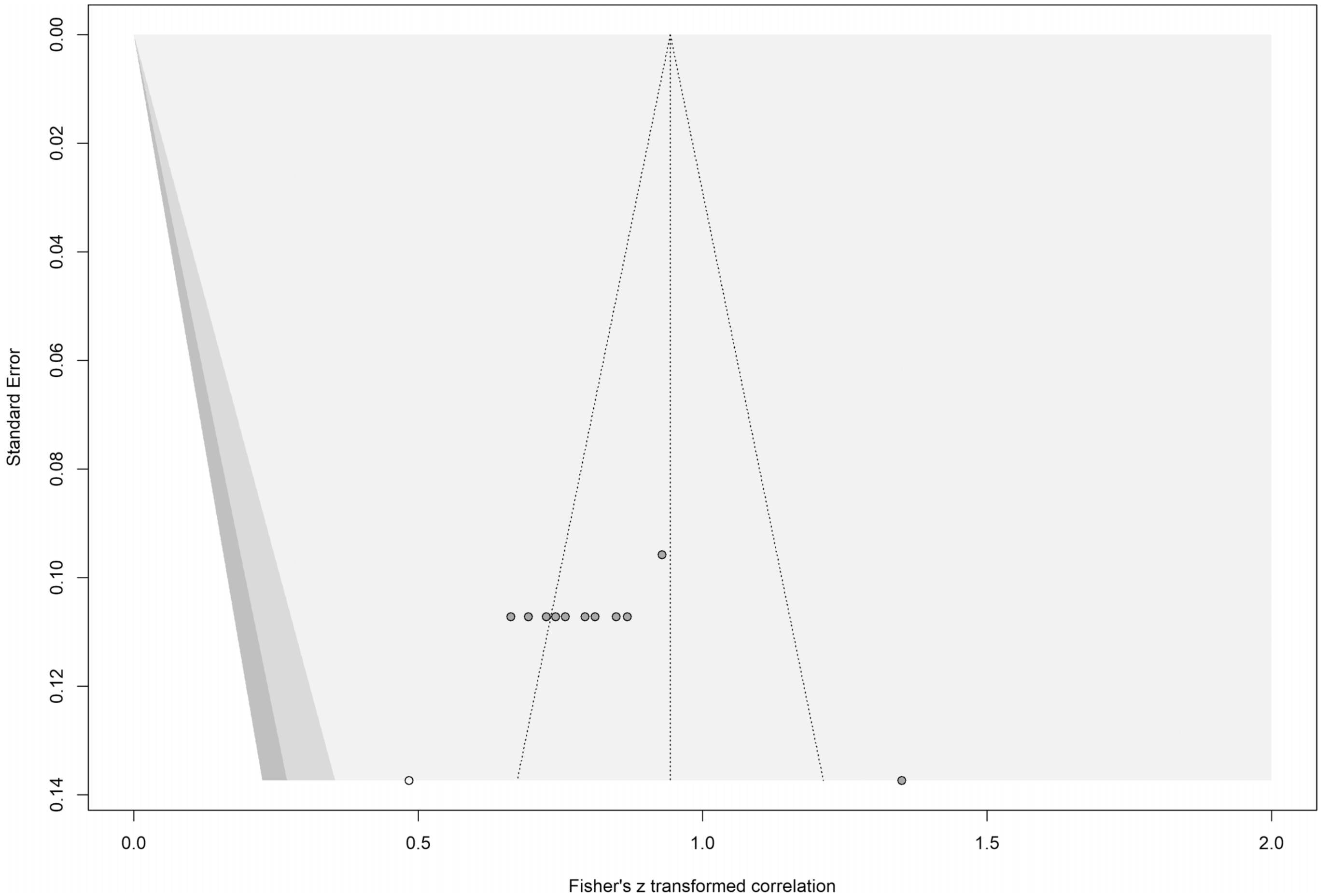

Fig. 5-B

Fig. 5 Funnel plot of (**Fig. 5-A**) intra-rater and (**Fig. 5-B**) inter-rater kappa coefficients of the Luo classification in the 3D-CT modality. 3D-CT = 3-dimensional computed tomography.

complicated fracture characteristics that are only discernible with the use of CT and MRI.

After performing meta-analysis, intrarater reliability of the conventional Schatzker classification (κ = 0.7218, 95% CI: 0.67-0.76) was more than that of the Hohl and Moore and conventional AO/OTA classifications, based on radiography. Acceptable reliability along with its simplicity made the Schatzker method the most common classification system used by trauma surgeons.

TABLE VI Landis and Koch Grading of Reliability

Kappa Value	Reliability Grading
<0	Poor agreement
0-0.2	Slight agreement
0.21-0.4	Fair agreement
0.41-0.6	Moderate agreement
0.61-0.8	Substantial agreement
0.8<	Excellent agreement

For the radiography modality, intrarater and inter-rater kappa coefficients of the Hohl and Moore classification were not acceptable. Despite its intriguing and useful concepts, it is mandatory to take necessary precautions before using this method.

It seems, after applying 2D-CT imaging, intrarater and inter-rater kappa coefficients of most of the classifications improved to more than 0.6, but none of them reached 0.7.

Considering 3D-CT images, the Luo classification showed the highest intrarater (κ = 0.8491, 95% CI: 0.83-0.86) and inter-rater (κ = 0.7718, 95% CI: 0.73-0.81) reliability, in comparison with the Schatzker (κ = 0.8129, 95% CI: 0.77-0.85; κ = 0.6573, 95% CI: 0.64-0.67; respectively) and AO/OTA (κ = 0.7041, 95% CI: 0.61-0.78; κ = 0.6218, 95% CI: 0.60-0.64; respectively) clasiification. In summary, Schatzker and Luo classifications are the most reliable methods based on radiography and 3D-CT scan, respectively (Fig. 6).

When comparing all mentioned kappa values together, among all 3 imaging modalities, the Luo classification system had the highest intrarater (κ = 0.8491) and inter-rater (κ = 0.7718)

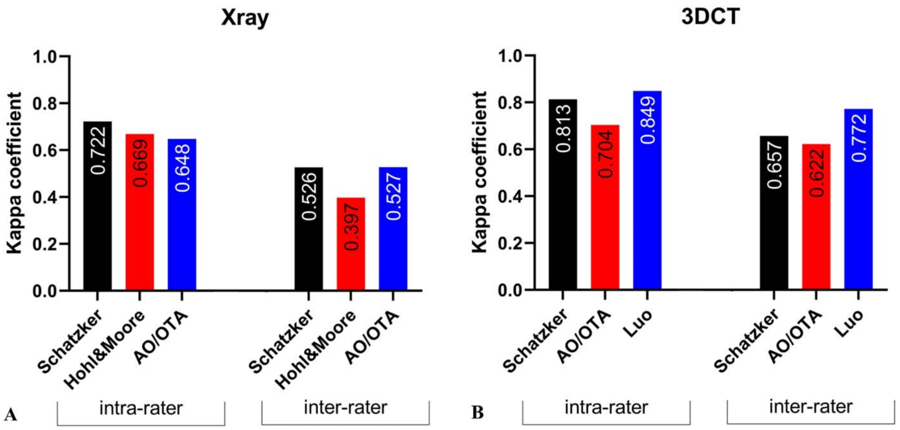

Fig. 6
Intrarater and inter-rater kappa coefficients in different TPF classifications. (**Fig. 6-A**) Radiography; (**Fig. 6-B**) 3D-CT. 3D-CT = 3-dimensional computed tomography.

kappa values. Nevertheless, there are some limitations. This classification system has not considered diaphyseal extension of the fracture, nor the concomitant injuries to the extensor mechanisms, tibial eminence, and fibular head. Also, it ignores the extent of soft-tissue involvement. Furthermore, newer potentially valuable concepts have been added to the literature including mechanism of injury[40] and main deformity direction[41]. Regarding publication bias after using the Egger's test and trim-and-fill method, Schatzker and Luo are still the most reliable classification systems in radiography and 3D-CT scan, respectively.

This study has some limitations. Meta-analysis was not performed for some methods, such as the revisited Schatzker classification system, due to lack of sufficient data. Improved reliability of a classification system may improve communication and research, but more research would need to be performed to address this, which if improved, changes the reliability of classifications, planning, or outcomes.

Future directions might include considering more advanced imaging like MRI and also adding some features to the Luo classification such as soft-tissue and associated central injuries.

Conclusion

Based on radiography, Schatzker showed a substantial intrarater and moderate inter-rater agreement. Nevertheless, comprehensive review of the history of TPF classifications' reliability from 1997 to 2023 showed that the 3-column classification proposed by Luo et al. was able to reach the highest degree and was the only classification with near-excellent inter-rater reliability.

Appendix

eA Supporting material provided by the authors is posted with the online version of this article as a data supplement at jbjs.org (http://links.lww.com/JBJSOA/A666). This content was not copy-edited or verified by JBJS. ■

Fardis Vosoughi, MD[1]
Iman Menbari Oskouie, MD, MPH[2]
Nazanin Rahimdoost, MD[3]
Amir Kasaeian, PhD[4,5,6]
Arash Sherafat Vaziri, MD[3]

[1]Department of orthopedics and trauma surgery, Tehran University of Medical Sciences, Tehran, Iran

[2]Urology Research Center, Tehran University of Medical Sciences, Tehran, Iran

[3]Center for Orthopedic Trans-Disciplinary Applied Research, Tehran University of Medical Sciences, Tehran, Iran

[4]Digestive Oncology Research Center, Digestive Diseases Research Institute, Shariati Hospital, Tehran University of Medical Sciences, Tehran, Iran

[5]Research Center for Chronic Inflammatory Diseases, Shariati Hospital, Tehran University of Medical Sciences, Tehran, Iran

[6]Clinical Research Development Unit, Shariati Hospital, Tehran University of Medical Sciences, Tehran, Iran

E-mail address for I. Menbari Oskouie: imanmenbary@gmail.com

References

1. Kfuri Júnior M, Fogagnolo F, Bitar RC, Freitas RL, Salim R, Paccola CAJ. Fraturas do planalto tibial. Rev Bras Ortop. 2009;44(6):468-74.

2. Jacofsky D, Haidukerwych G, Scott W. Tibia plateau fractures. In: Scott WN Insall & Scott Surgery of the Knee. Philadelphia, PA: Churchill Livingstone; 2006:1133-46.

3. Charalambous C, Tryfonidis M, Alvi F, Moran M, Fang C, Samarji R, Hirst P. Inter- and intra-observer variation of the Schatzker and AO/OTA classifications of tibial plateau fractures and a proposal of a new classification system. Ann R Coll Surg Engl. 2007;89(4):400-4.

4. Hohl M. Tibial condylar fractures. J Bone Joint Surg Am. 1967;49(7):1455-67.

5. Evarts CM. Surgery of the Musculoskeletal System. London, United Kingdom: Churchill Livingstone; 1990.

6. Moore TM. Fracture-dislocation of the knee. Clin Orthop Relat Res. 1981;156: 128-40.

7. Müller ME, Nazarian S, Koch P, Schatzker J. The Comprehensive Classification of Fractures of Long Bones. Berlin, Germany: Springer Science & Business Media; 2012.

8. Schatzker J, Mcbroom R, Bruce D. The tibial plateau fracture: the Toronto experience 1968-1975. Clin Orthop Relat Res. 1979;138:94-104.

9. Wahlquist M, Iaguilli N, Ebraheim N, Levine J. Medial tibial plateau fractures: a new classification system. J Trauma. 2007;63(6):1418-21.

10. Marsh J, Slongo TF, Agel J, Broderick JS, Creevey W, DeCoster TA, Prokuski L, Sirkin MS, Ziran B, Henley B, Audigé L. Fracture and dislocation classification compendium-2007: Orthopaedic Trauma Association Classification, Database and Outcomes Committee. J Orthop Trauma. 2007;21(10):s1-s6.

11. Martínez-Rondanelli A, Escobar-González SS, Henao-Alzate A, Martínez-Cano JP. Reliability of a four-column classification for tibial plateau fractures. Int Orthop. 2017;41(9):1881-6.

12. Chang SM, Hu SJ, Zhang YQ, Yao MW, Ma Z, Wang X, Dargel J, Eysel P. A surgical protocol for bicondylar four-quadrant tibial plateau fractures. Int Orthop. 2014;38(12):2559-64.

13. Molenaars RJ, Mellema JJ, Doornberg JN, Kloen P. Tibial plateau fracture characteristics: computed tomography mapping of lateral, medial, and bicondylar fractures. J Bone Joint Surg Am. 2015;97(18):1512-20.

14. Hu YL, Ye FG, Ji AY, Qiao GX, Liu HF. Three-dimensional computed tomography imaging increases the reliability of classification systems for tibial plateau fractures. Injury. 2009;40(12):1282-5.

15. Doornberg JN, Rademakers MV, van den Bekerom MP, Kerkhoffs GM, Ahn J, Steller EP, Kloen P. Two-dimensional and three-dimensional computed tomography for the classification and characterisation of tibial plateau fractures. Injury. 2011; 42(12):1416-25.

16. Yang G, Zhai Q, Zhu Y, Sun H, Putnis S, Luo C. The incidence of posterior tibial plateau fracture: an investigation of 525 fractures by using a CT-based classification system. Arch Orthop Trauma Surg. 2013;133(7):929-34.

17. Aghaghazvini L, Tahmasebi MN, Gerami R, Vaziri AS, Rasuli B, Tahami M, Vosoughi F. Sonography: a sensitive and specific method for detecting trochlear cartilage pathologies. J Ultrasound. 2020;23(3):259-63.

18. Luo CF, Sun H, Zhang B, Zeng BF. Three-column fixation for complex tibial plateau fractures. J Orthop Trauma. 2010;24(11):683-92.

19. Tunçez M, Akan I, Seyfettinoglu F, Çetin Tunçez H, Dirim Mete B, Kazımoğlu C. Is it necessary to add soft tissue injury to the classification in tibial plateau fracture management? Cureus. 2022;14(2):e22236.

20. Maripuri SN, Rao P, Manoj-Thomas A, Mohanty K. The classification systems for tibial plateau fractures: how reliable are they? Injury. 2008;39(10):1216-21.

21. Walton NP, Harish S, Roberts C, Blundell C. AO or Schatzker? How reliable is classification of tibial plateau fractures? Arch Orthop Trauma Surg. 2003;123(8):396-8.

22. Brunner A, Horisberger M, Ulmar B, Hoffmann A, Babst R. Classification systems for tibial plateau fractures; does computed tomography scanning improve their reliability? Injury. 2010;41(2):173-8.

23. Ahmad S, Arora R, Aggarwal AN, Ul-Haq R, Khan K, Tandon A. Comparison between reliability of Schatzker's classification and CT based four quadrant classification for tibial plateau fractures. J Clin Orthop Trauma. 2022;32:101986.

24. Yacoubian SV, Nevins RT, Sallis JG, Potter HG, Lorich DG. Impact of MRI on treatment plan and fracture classification of tibial plateau fractures. J Orthop Trauma. 2002;16(9):632-7.

25. Rethlefsen ML, Kirtley S, Waffenschmidt S, Ayala AP, Moher D, Page MJ, Koffel JB, Blunt H, Brigham T, Chang S, Clark J, Conway A, Couban R, de Kock S, Farrah K, Fehrmann P, Foster M, Fowler SA, Glanville J, Harris E, Hoffecker L, Isojarvi J, Kaunelis D, Ket H, Levay P, Lyon J, McGowan J, Murad MH, Nicholson J, Pannabecker V, Paynter R, Pinotti R, Ross-White A, Sampson M, Shields T, Stevens A, Sutton A, Weinfurter E, Wright K, Young S. PRISMA-S: an extension to the PRISMA statement for reporting literature searches in systematic reviews. Syst Rev. 2021;10(1):39-19.

26. McGowan J, Sampson M, Salzwedel DM, Cogo E, Foerster V, Lefebvre C. PRESS peer review of electronic search strategies: 2015 guideline statement. J Clin Epidemiol. 2016;75:40-6.

27. Hupe M. EndNote X9. J Electron Resour Med Libr. 2019;16(3-4):117-9.

28. Ouzzani M, Hammady H, Fedorowicz Z, Elmagarmid A. Rayyan—a web and mobile app for systematic reviews. Syst Rev. 2016;5:210-0.

29. Vosoughi F, Makuku R, Tantuoyir MM, Yousefi F, Shobeiri P, Karimi A, Alilou S, LaPorte R, Tilves C, Nabian MH, Yekaninejad MS. A systematic review and meta-analysis of the epidemiological characteristics of COVID-19 in children. BMC Pediatr. 2022;22(1):613.

30. Valizadeh A, Moassefi M, Nakhostin-Ansari A, Oskoie IM, Some'eh SH, Aghajani F, Torbati M, Maleki Ghorbani Z, Aghajani R, Hosseini Asl SH, Mirzamohammadi A, Ghafouri M, Faghani S, Memari AH. Accuracy of machine learning algorithms for the diagnosis of autism spectrum disorder based on cerebral sMRI, rs-fMRI, and EEG: protocols for three systematic reviews and meta-analyses. medRxiv. 2021. https://doi.org/10.1101/2021.06.29.21254249.

31. Institute JB. Critical Appraisal Tools for Use in JBI Systematic Reviews. Adelaide, AU: JBI; 2017.

32. Moola S, Munn Z, Tufanaru C, Aromataris E, Sears K, Sfetcu R, Currie M, Lisy K, Qureshi R, Mattis P, Mu P. Chapter 7: systematic reviews of etiology and risk. In: Joanna Briggs Institute Reviewer's Manual. Vol 5. Adelaide, Australia: The Joanna Briggs Institute; 2017.

33. Ma LL, Wang YY, Yang ZH, Huang D, Weng H, Zeng XT. Methodological quality (risk of bias) assessment tools for primary and secondary medical studies: what are they and which is better? Mil Med Res. 2020;7:7-11.

34. R Core Team R. R: a language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing 2013.

35. Furuya-Kanamori L, Kostoulas P, Doi SA. A new method for synthesizing test accuracy data outperformed the bivariate method. J Clin Epidemiol. 2021;132:51-8.

36. Higgins JP, Thompson SG. Quantifying heterogeneity in a meta-analysis. Stat Med. 2002;21(11):1539-58.

37. Audigé L, Bhandari M, Hanson B, Kellam J. A concept for the validation of fracture classifications. J Orthop Trauma. 2005;19(6):401-6.

38. Millar SC, Arnold JB, Thewlis D, Fraysse F, Solomon LB. A systematic literature review of tibial plateau fractures: what classifications are used and how reliable and useful are they? Injury. 2018;49(3):473-90.

39. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics. 1977;33(1):159-74.

40. Xie X, Zhan Y, Wang Y, Lucas JF, Zhang Y, Luo C. Comparative analysis of mechanism-associated 3-dimensional tibial plateau fracture patterns. J Bone Joint Surg Am. 2020;102(5):410-8.

41. Boluda-Mengod J, Guimerà-García V, Olías-López B, Renovell-Ferrer P, Carrera I, Herrera-Pérez M, Pais-Brito JL. A proposal of a new algorithm for decision-making approaches in open reduction and internal fixation of complex tibial plateau fractures–SOTA algorithm (Spanish Orthopaedic Trauma Association). Injury. 2021;52(suppl 4):S87-s98.

42. Chan PS, Klimkiewicz JJ, Luchetti WT, Esterhai JL, Kneeland JB, Dalinka MK, Heppenstall RB. Impact of CT scan on treatment plan and fracture classification of tibial plateau fractures. J Orthop Trauma. 1997;11(7):484-9.

43. Gicquel T, Najihi N, Vendeuvre T, Teyssedou S, Gayet LE, Huten D. Tibial plateau fractures: reproducibility of three classifications (Schatzker, AO, Duparc) and a revised Duparc classification. Orthop Traumatol Surg Res. 2013;99(7):805-16.

44. Martin J, Marsh JL, Nepola JV, Dirschl DR, Hurwitz S, DeCoster TA. Radiographic fracture assessments: which ones can we reliably make? J Orthop Trauma. 2000; 14(6):379-85.

45. Mellema JJ, Doornberg JN, Molenaars RJ, Ring D, Kloen P, Traumaplatform Study Collaborative & Science of Variation Group. Interobserver reliability of the Schatzker and Luo classification systems for tibial plateau fractures. Injury. 2016;47(4):944-9.

46. Patange Subba Rao SP, Lewis J, Haddad Z, Paringe V, Mohanty K. Three-column classification and Schatzker classification: a three- and two-dimensional computed tomography characterisation and analysis of tibial plateau fractures. Eur J Orthop Surg Traumatol. 2014;24(7):1263-70.

47. te Stroet MA, Holla M, Biert J, van Kampen A. The value of a CT scan compared to plain radiographs for the classification and treatment plan in tibial plateau fractures. Emerg Radiol. 2011;18(4):279-83.

48. Wennergren D, Ekholm C, Sundfeldt M, Karlsson J, Bhandari M, Möller M. High reliability in classification of tibia fractures in the Swedish Fracture Register. Injury. 2016;47(2):478-82.

49. Zhu Y, Hu CF, Yang G, Cheng D, Luo CF. Inter-observer reliability assessment of the Schatzker, AO/OTA and three-column classification of tibial plateau fractures. J Trauma Manag Outcomes. 2013;7(1):7.

50. Zhu Y, Yang G, Luo CF, Smith WR, Hu CF, Gao H, Zhong B, Zeng BF. Computed tomography-based three-column classification in tibial plateau fractures: introduction of its utility and assessment of its reproducibility. J Trauma Acute Care Surg. 2012;73(3):731-7.

51. Hoekstra H, Kempenaers K, Nijs S. A revised 3-column classification approach for the surgical planning of extended lateral tibial plateau fractures. Eur J Trauma Emerg Surg. 2017;43(5):637-43.

52. Albuquerque RP, Giordano V, Pallottino A, Sassine T, Canedo R, Pina J, do Amaral NP. Analysis of the reproducibility of tibial plateau fractures' classification. Rev Bras Ortop. 2009;44(3):225-9.

53. Crijns TJ, Mellema JJ, Özkan S, Ring D, Chen NC, Science of Variation Group. Classification of tibial plateau fractures using 3DCT with and without subtraction of unfractured bones. Injury. 2020;51(11):2686-91.

54. Alencar JBD, Souza CJDD, Coelho PRR, Freitas PGD, Lopes MBG, Pinto FJ. Inter- and intra-observer reliability of Schatzker, Ao, and Luo classifications for tibial plateau fractures. Acta Ortop Bras. 2020;28(5):216-20.

55. Masouros PT, Mitrogiannis G, Antoniou G, Chatzidaki C, Kourtzis D, Garnavos C. Interobserver reliability of schatzker, AO foundation-orthopaedic trauma association, and Luo classifications for tibial plateau fractures: does three-dimensional CT improve outcomes? Cureus. 2022;14(2):e22227.

56. Khan MA, Pal S, Ahmed SK, Chinoy MA. Inter-observer variation of the Schatzker and Khan classification of Tibial plateau fractures: cohort study. J Pak Med Assoc. 2021;71(Suppl 5):S51-4.

57. Yao P, Liu Y, Shan L, Zhao Y, Wang D, He Y, Wang H, Zeeshan Waheed M, Lu T, Zhou J. Intra- and inter-observer reliability assessment of widely used classifications and the "Ten-segment classification" of tibial plateau fractures. Knee. 2022;35:149-56.

58. Taşkesen A, Demirkale İ, Okkaoğlu MC, Özdemir M, Bilgili MG, Altay M. Intraobserver and interobserver reliability assessment of tibial plateau fracture classification systems. Eklem Hastalik Cerrahisi. 2017;28(3):177-81.

59. Zhang BB, Sun H, Zhan Y, He QF, Zhu Y, Wang YK, Luo CF. Reliability and repeatability of tibial plateau fracture assessment with an injury mechanism-based concept. Bone Joint Res. 2019;8(8):357-66.

60. Dust T, Hartel MJ, Henneberg JE, Korthaus A, Ballhause TM, Keller J, Ohlmeier M, Maas KJ, Frosch KH, Krause M. The influence of 3D printing on inter-and intrarater reliability on the classification of tibial plateau fractures. Eur J Trauma Emerg Surg. 2022;49:189-99.

61. Anwar A, Zhang Y, Zhao Z, Gao Y, Sha L, Lv D, Zhang Z, Lv G, Zhang Y, Nazir MU, Qasim W, Wang Y. Two column classification of tibial plateau fractures; description, clinical application and reliability. Injury. 2019;50(6):1247-55.

62. Millán-Billi A, Gómez-Masdeu M, Ramírez-Bermejo E, Ibañez M, Gelber PE. What is the most reproducible classification system to assess tibial plateau fractures? Int Orthop. 2017;41(6):1251-6.

63. Mansur H, Corrêa VLB, Abdo B, Ramos LS, Castiglia MT. Evaluation of the reproducibility of the schatzker classification reviewed by Kfuri for tibial plateau fractures. Rev Bras Ortop. 2022;57(3):502-10.

64. Huitema JM, van der Gaast N, Brouwers L, Jaarsma RL, Doornberg JN, Edwards MJ, Hermans E. Are 3D printed models of tibial plateau fractures a useful addition to understanding fractures for junior surgeons? Clin Orthop Relat Res. 2022;10:1097.

65. Pätzold R, Friederichs J, von Rüden C, Panzer S, Bühren V, Augat P. The pivotal role of the coronal fracture line for a new three-dimensional CT-based fracture classification of bicondylar proximal tibial fractures. Injury. 2017;48(10):2214-20.

66. Castiglia MT, Nogueira-Barbosa MH, Messias AMV, Salim R, Fogagnolo F, Schatzker J, Kfuri M. The impact of computed tomography on decision making in tibial plateau fractures. J Knee Surg. 2018;31(10):1007-14.