ORIGINAL ARTICLE

Comparison of Raters' Reliability of Two Methods for Measuring Tooth Preparations Undercuts using Professional Computer-Aided Design Software

Abdulrahman Alhaddad, Thamer Marghalani

Department of Oral and Maxillofacial Prosthodontics, Faculty of Dentistry, King Abdulaziz University, Jeddah, Saudi Arabia

Abstract

Background: Tooth preparation for fixed dental prostheses requires achieving adequate biological, functional, and aesthetic requirements. One critical aspect of preparation is avoiding undercuts, which can compromise the prosthesis' fit, retention, and resistance. Detection of undercuts is traditionally done visually, but advancements in computer-aided design (CAD) software provide new methods for assessing undercuts. This study compares the reliability of two professional CAD software – Fusion 360 and ExoCad – in detecting undercuts in tooth preparations. **Materials and Methods:** Seventy-two tooth preparations from 26 dental stone casts were scanned and digitized into 3D models. Two experienced prosthodontists analyzed undercuts in the 3D models using Fusion 360 and ExoCad. Interclass correlation coefficient (ICC) values were calculated to evaluate intra- and inter-rater reliability between the software and operators. **Results:** The ICC analysis indicated excellent intra-rater reliability for one operator using both software, while the other operator showed poor to fair reliability. Inter-rater reliability was excellent for ExoCad in some instances but varied significantly across locations and software. The overall presence of undercuts ranged from 85.3% to 98.7%. **Conclusion:** Variations in undercut detection between the software were operator-dependent, highlighting the importance of operator skill in aligning the insertion path. Future studies should explore multi-unit preparations and unified CAD algorithms for undercut detection.

Key words: Computer-aided design software, tooth preparation, undercut

INTRODUCTION

Many reasons cause missing teeth and lost tooth structures, including dental caries, trauma, periodontal diseases, and other reasons. Dental prostheses or restorations can replace total or partial loss of tooth structures. Dental prostheses, like crowns and fixed dental prostheses, are made by preparing tooth surfaces with precision to receive coronal coverages necessary to replace missing teeth or tooth parts. Failures of fixed prosthodontics are mainly caused by inadequate tooth preparation, which can be in the biological aspects such as caries, periodontal, and endodontic complications or/and in the aesthetic aspect like over-contouring.[1] Inadequate tooth preparation can lead to mechanical failures such as loss of retention and resistance form, which is the main cause of loosening and dislodgment of the prosthesis in the long run.[1] One of the important requirements for the fit of fixed prosthodontics is to have adequate tooth preparation. Tooth preparation requires adequate biological, functional, and aesthetic requirements.[2] Having parallel opposing walls in tooth preparation will conserve tooth structure and enhance retention and resistance forms.[3] Achieving parallel walls is hard, especially in the clinical setting, without causing undercuts. A minimal 12° taper ensures an undercut-free preparation. The inclination of walls in the posterior teeth may be affected by poor accessibility,

Address for correspondence:
Abdulrahman Alhaddad, Department of Oral and Maxillofacial Prosthodontics, Faculty of Dentistry, King Abdulaziz University, Jeddah, Saudi Arabia.
E-mail: aalhaddad@kau.edu.sa

Received: 14-11-2025
Revised: 17-12-2025
Accepted: 27-12-2025

which might force the clinician to a compromised taper and inclination. Furthermore, a degree of convergence is needed to recover for the possible manufacturing processes and to allow estimation of the preparation walls and prevention of undercuts, compensate for manufacturing errors, and allow a more suitable seat during cementation.[4] Retention and resistance are the most important concepts in crown preparation. "Retention" is defined as the ability of a crown to resist pulling forces acting along its path of insertion (POI). At the same time, "resistance" is the ability of a crown to resist forces acting to dislodge it in a direction other than along the POI.

The operator can inadvertently create undercuts when trying to achieve the necessary taper or convergence angles to fit the prosthesis while striving to have parallel walls of the preparation surfaces necessary for achieving prosthesis retention. A balanced approach is required to reach adequate tooth preparation. Nevertheless, undercuts do occur. Detection of their presence requires visual checking by the operator inside the patient's mouth, with the limited accessibility and restrictions present in the mouth, before taking the final impressions for the tooth preparations. Once taken, the impressions can be evaluated easily for undercuts using the visual method. So far, experienced clinicians can visually determine undercuts presence or occurrences. Only a little literature is present on evaluating undercut using computer software. Computer software was used to determine the presence of undercuts using professional dental software like ExoCad. This article introduces and evaluates a new method to determine the occurrences of undercuts using professional computer-aided design (CAD) software Fusion360 compared to ExoCad. Each software utilizes the operator to determine a certain insertion path before mapping the undercut's location. Reliability analysis is usually used to determine the levels of agreement between operators in determining the presence of undercuts utilizing both software (Ref). Reliability analysis was done for two experienced prosthodontic raters who assign a POI in the software then each software will map the presence and location of the undercut.

All studies mentioned and measured ideal or achievable convergence angles from preparations made in near-ideal conditions on typodonts. None were measured on models generated from actual patients in real clinical situations. Only one study measured the undercuts and POI on typodonts for students in an examination setting.

This study compares two innovative methods of measuring undercuts that have not been mentioned or done in the literature before. The two methods have been validated by measuring the reliability analysis interclass correlation coefficient (ICC) between two experienced prosthodontist raters.

Objectives of the study

This study is different from the other studies in that it will measure the reliability (agreement) in detecting the occurrence of undercuts in clinically produced patients' tooth preparations utilizing the 3D digital scanned models of the prepared tooth using dental CAD/computer-aided manufacturing scanners and two professional CAD software operated by two experienced prosthodontist raters.

Hypothesis

The null hypothesis was that there was no difference between the raters and the software in the occurrences of undercuts measured using CAD software in 3D models of tooth preparations obtained from clinical patients' cases.

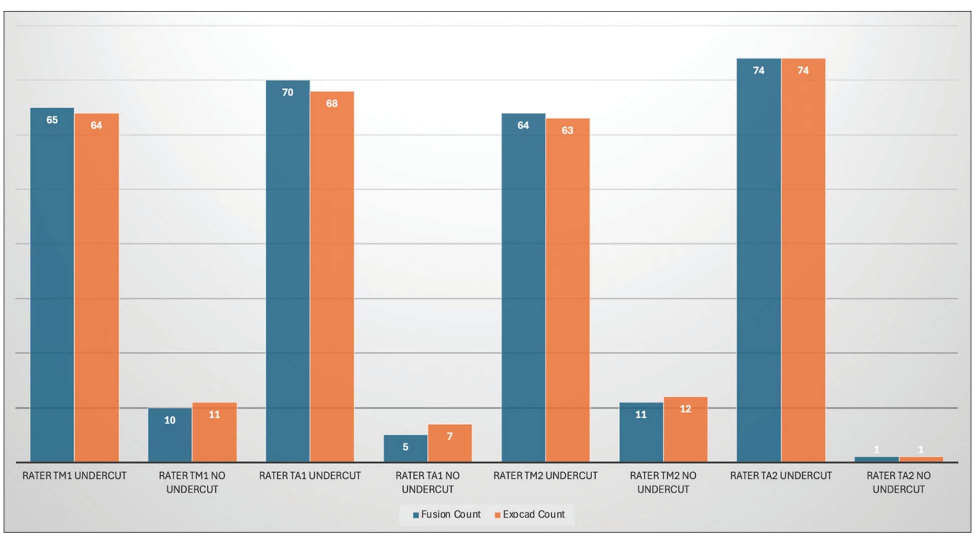

Figure 1: Undercut presence counts by raters and software

MATERIALS AND METHODS

Twenty-six dental stone casts were collected randomly from dental schools, dental laboratories, and general and private practices in Jeddah, Saudi Arabia, for patients treated by students, general practitioners, and prosthodontic specialists. Damaged casts were excluded from the study. The sample comprised 72 preparations obtained from the 26 dental stone models. The stone casts were die-trimmed, ditched, 3D scanned, digitized, and saved in STL format using a 3D cast scanner (Arctica, KAVO Dental) and specialized dental software (Exocad, Exocad GmbH). The scanned models' STL files were then imported to the (MeshMixer, AutoDesk Inc.,), where the mesh was segmented, cleaned, made as solid objects, and then reduced facet size to the required specifications of the Fusion 360 software (Fusion360, AutoDesk Inc.,), which is 20,000 facets. A video that shows how the STL file was processed in MeshMixer is presented in the attached video 1. Then, the mesh was imported into Fusion 360 and ExoCad software, where two operators measure it for the presence of any undercuts on the preparations regardless of its location. Then, to ensure a high-level intra-examiner reliability, only one person measured the same tooth angles and undercut twice within 2 months. The location of the undercut is noted if it was present, whether in mesial (M), distal (D), buccal (B), or lingual (L) surfaces of the prepared tooth. A video showing how the mesh file was processed and angles were measured in Fusion360 and ExoCad is presented in the attached videos 2 and 3, respectively. Descriptive statistics, ICC, and Pearson Chi-square statistical analyses of the data were performed using the Statistical Package for the Social Sciences 29 software package, IBM. ICC estimates and their 95% confidence intervals were calculated, once based on mean-rating (k = 1), absolute agreement, and a two-way mixed-effect model to check for the agreement between the software by the same rater. Moreover, in the 2nd time, based on mean-rating (k = 2), absolute agreement, a 2-way mixed-effect model was used to check the agreement between the two raters using the same software. Cicchetti provided the following interpretation guidelines for ICC values:

ICC < 0.40 = Poor, ICC 0.40–0.59 = Fair, ICC 0.60–0.74 = Good, ICC ≥ 0.75 = Excellent.[5] According to McGraw and Wong[6] and Shrout and Fliess,[7] "values <0.5 are indicative of poor reliability, values between 0.5 and 0.75 indicate moderate reliability, values between 0.75 and 0.9 indicate good reliability, and values >0.90 indicate excellent reliability."

RESULTS

Detailed ICC, 95% confidence intervals, significance, and levels of inter-rater reliability are shown in Tables 1-5. For intra-rater reliability, each rater was compared to the

Table 1: Interclass correlation coefficient showing intra-rater reliability (within each rater) for the presence of undercut.

Undercut type/Location	Software (s)	Rater (s)	Rating time (1st or 2nd)	Interpretation of intraclass correlation coefficient	Reliability	Intraclass correlation coefficient: Intraclass correlation[b]	95% confidence interval: Lower bound	95% confidence interval: Upper bound	F-test with true value-0: Value	df1	df2	Sig
Presence	ExoCad	TM	1st and 2nd	Good	Single Measures	0.644[a]	0.489	0.759	4.573	74	74	<0.001
				Excellent	Average Measures	0.783[c]	0.657	0.863	4.573	74	74	<0.001
Presence	Fusion 360	TM	1st and 2nd	Good	Single Measures	0.616[a]	0.453	0.739	4.168	74	74	<0.001
				Excellent	Average Measures	0.762[c]	0.623	0.850	4.168	74	74	<0.001
Presence	Fusion 360	TA	1st and 2nd	Poor	Single Measures	0.321[a]	0.109	0.507	1.986	74	74	0.002
				Fair	Average Measures	0.486[c]	0.196	0.673	1.986	74	74	0.002
Presence	ExoCad	TA	1st and 2nd	Poor	Single Measures	0.234[a]	0.021	0.431	1.657	74	74	0.016
				Poor	Average Measures	0.380[c]	0.040	0.603	1.657	74	74	0.016

[a]The estimator is the same, whether the interaction effect is present or not. [b]Type A intraclass correlation coefficients using an absolute agreement definition. [c]This estimate is computed assuming the interaction effect is absent, because it is not estimable otherwise.

Table 2: Interclass correlation coefficient showing intra -rater reliability (within each rater) for locations of undercut

Interpretation of intraclass correlation coefficient						Intraclass correlation coefficient						
Undercut type/location	Software (s)	Rater (s)	Rating time (1st or 2nd)	Reliability		Intraclass correlation[b]	95% confidence interval		F-test with true value-0			
							Lower bound	Upper bound	Value	df1	df2	Sig
Mesial	Fusion 360	TM	1st and 2nd	Fair	Single Measures	0.412[a]	0.206	0.583	2.404	74	74	<.001
				Fair	Average Measures	0.583[c]	0.342	0.736	2.404	74	74	<0.001
Distal	Fusion 360	TM	1st and 2nd	Poor	Single Measures	0.362[a]	0.147	0.544	2.120	74	74	<0.001
				Fair	Average Measures	0.532[c]	0.256	0.705	2.120	74	74	<0.001
Buccal	Fusion 360	TM	1st and 2nd	Fair	Single Measures	0.475[a]	0.281	0.632	2.911	74	74	<0.001
				Good	Average Measures	0.644[c]	0.438	0.775	2.911	74	74	<0.001
Lingual	Fusion 360	TM	1st and 2nd	Poor	Single Measures	0.337[a]	0.127	0.520	2.092	74	74	<.001
				Fair	Average Measures	0.504[c]	0.226	0.684	2.092	74	74	<0.001
Mesial	ExoCad	TM	1st and 2nd	Fair	Single Measures	0.567[a]	0.391	0.703	3.587	74	74	<0.001
				Good	Average Measures	0.724[c]	0.562	0.826	3.587	74	74	<0.001
Distal	ExoCad	TM	1st and 2nd	Good	Single Measures	0.735[a]	0.611	0.824	6.501	74	74	<0.001
				Excellent	Average Measures	0.847[c]	0.758	0.903	6.501	74	74	<0.001
Buccal	ExoCad	TM	1st and 2nd	Poor	Single Measures	0.344[a]	0.133	0.526	2.081	74	74	<0.001
				Fair	Average Measures	0.512[c]	0.235	0.690	2.081	74	74	<0.001
Lingual	ExoCad	TM	1st and 2nd	Fair	Single Measures	0.573[a]	0.398	0.707	3.647	74	74	<0.001
				Good	Average Measures	0.728[c]	0.569	0.829	3.647	74	74	<0.001
Mesial	Fusion 360	TA	1st and 2nd	Good	Single Measures	0.666[a]	0.517	0.775	4.929	74	74	<0.001
				Excellent	Average Measures	0.799[c]	0.682	0.873	4.929	74	74	<0.001
Distal	Fusion 360	TA	1st and 2nd	Good	Single Measures	0.656[a]	0.506	0.767	4.875	74	74	<0.001
				Excellent	Average Measures	0.792[c]	0.672	0.868	4.875	74	74	<0.001
Buccal	Fusion 360	TA	1st and 2nd	Good	Single Measures	0.729[a]	0.603	0.819	6.411	74	74	<0.001
				Excellent	Average Measures	0.843[c]	0.752	0.901	6.411	74	74	<0.001
Lingual	Fusion 360	TA	1st and 2nd	Good	Single Measures	0.729[a]	0.603	0.819	06.411	74	74	<0.001
				Excellent	Average Measures	0.843[c]	0.752	0.901	6.411	74	74	<0.001
Mesial	ExoCad	TA	1st and 2nd	Poor	Single Measures	0.296[a]	0.073	0.490	1.828	74	74	0.005
				Fair	Average Measures	0.456[c]	0.136	0.657	1.828	74	74	0.005
Distal	ExoCad	TA	1st and 2nd	Poor	Single Measures	0.297[a]	0.085	0.485	1.914	74	74	0.003
				Fair	Average Measures	0.457[c]	0.156	0.653	1.914	74	74	0.003
Buccal	ExoCad	TA	1st and 2nd	Poor	Single Measures	0.279[a]	0.056	0.476	1.768	74	74	0.008
				Fair	Average Measures	0.437[c]	0.106	0.645	1.768	74	74	0.008
Lingual	ExoCad	TA	1st and 2nd	Poor	Single Measures	0.391[a]	0.175	0.570	2.477	74	74	<0.001
				Fair	Average Measures	0.562[c]	0.297	0.726	2.477	74	74	<0.001

[a]The estimator is the same, whether the interaction effect is present or not. [b]Type A intraclass correlation coefficients using an absolute agreement definition. [c]This estimate is computed assuming the interaction effect is absent, because it is not estimable otherwise.

Table 3: Interclass correlation coefficient showing inter-rater reliability within and between the software for the presence of undercut

Interpretation of intraclass correlation coefficient						Intraclass correlation coefficient						
Undercut type/location	Software (s)	Rater (s)	Rating time (1st or 2nd)		Reliability	Intraclass correlation[b]	95% confidence interval		F-test with true value-0			
							Lower bound	Upper bound	Value	df1	df2	Sig
Presence	ExoCad	TM	1st	Good	Single Measures	0.627[a]	0.468	0.747	4.438	74	74	<0.001
		TA	1st	Excellent	Average Measures	0.771[c]	0.638	0.855	4.438	74	74	<0.001
Presence	Fusion 360	TM	1st	Good	Single Measures	0.637[a]	0.478	0.755	4.714	74	74	<0.001
		TA	1st	Excellent	Average Measures	0.778[c]	0.647	0.861	4.714	74	74	<0.001
Presence	Fusion 360	TA	2nd	Poor	Single Measures	−0.025[a]	−0.223	0.183	0.945	74	74	0.596
		TM	2nd	Poor	Average Measures	−0.052[c]	−0.573	0.310	0.945	74	74	0.596
Presence	ExoCad	TA	2nd	Good	Single Measures	0.644[a]	0.489	0.759	4.573	74	74	<0.001
		TM	2nd	Excellent	Average Measures	0.783[c]	0.657	0.863	4.573	74	74	<0.001
Presence	ExoCad Fusion 360	TM	1st	Good	Single Measures	0.644[a]	0.489	0.759	4.573	74	74	<0.001
		TA	1st	Excellent	Average Measures	0.783[c]	0.657	0.863	4.573	74	74	<0.001
Presence	ExoCad Fusion 360	TA	2nd	Poor	Single Measures	−0.025[a]	−0.223	0.183	0.945	74	74	0.596
		TM	2nd	Poor	Average Measures	−0.052[c]	−0.573	0.310	0.945	74	74	0.596
Presence	ExoCad	TM	2nd	Fair	Single Measures	0.466[a]	0.272	0.625	2.791	74	74	<0.001
		TA	1st	Good	Average Measures	0.636[c]	0.427	0.769	2.791	74	74	<0.001
Presence	Fusion 360	TM	2nd	Fair	Single Measures	0.453[a]	0.256	0.614	2.738	74	74	<0.001
		TA	1st	Good	Average Measures	0.623[c]	0.407	0.761	2.738	74	74	<0.001
Presence	Fusion 360	TA	2nd	Poor	Single Measures	0.163[a]	−0.044	0.364	1.438	74	74	0.060
		TM	1st	Poor	Average Measures	0.281[c]	−0.093	0.533	1.438	74	74	0.060
Presence	ExoCad	TA	2nd	Poor	Single Measures	0.163[a]	−0.044	0.364	1.438	74	74	0.060
	Fusion 360	TM	1st	Poor	Average Measures	0.281[c]	−0.093	0.533	1.438	74	74	0.060
Presence	Fusion	TA	2nd	Good	Single Measures	0.644[a]	0.489	0.759	4.573	74	74	<0.001
	ExoCad	TM	1st	Excellent	Average Measures	0.783[c]	0.657	0.863	4.573	74	74	<0.001
Presence	ExoCad	TA	2nd	Good	Single Measures	0.644[a]	0.489	0.759	4.573	74	74	<0.001
	Fusion 360	TM	1st	Excellent	Average Measures	0.783[c]	0.657	0.863	4.573	74	74	<0.001

[a]The estimator is the same, whether the interaction effect is present or not. [b]Type A intraclass correlation coefficients using an absolute agreement definition. [c]This estimate is computed assuming the interaction effect is absent, because it is not estimable otherwise.

Table 4: Interclass correlation coefficient showing inter-rater reliability within the software and between the 1st time and 2nd time of measurement

Interpretation of intraclass correlation coefficient						Intraclass correlation coefficient						
Undercut type/location	Software (s)	Rater (s)	Rating time (1st or 2nd)		Reliability	Intraclass correlation[b]	95% confidence interval		F-test with true value-0			
							Lower bound	Upper bound	Value	df1	df2	Sig
Mesial	Fusion 360	TM	1st	Poor	Single Measures	0.255[a]	0.040	0.450	1.819	74	74	0.005
		TA	1st	Fair	Average Measures	0.406[c]	0.076	0.620	1.819	74	74	0.005
Distal	Fusion 360	TM	1st	Poor	Single Measures	0.327[a]	0.116	0.512	2.026	74	74	0.001
		TA	1st	Fair	Average Measures	0.493[c]	0.208	0.677	2.026	74	74	0.001
Buccal	Fusion 360	TM	1st	Poor	Single Measures	0.114[a]	−0.093	0.318	1.284	74	74	0.142
		TA	1st	Poor	Average Measures	0.204[c]	−0.204	0.482	1.284	74	74	0.142
Lingual	Fusion 360	TM	1st	Poor	Single Measures	0.240[a]	0.027	0.436	1.759	74	74	0.008
		TA	1st	Poor	Average Measures	0.387[c]	0.053	0.607	1.759	74	74	0.008
Mesial	ExoCad	TM	1st	Excellent	Single Measures	0.836[a]	0.752	0.893	11.406	74	74	<0.001
		TA	1st	Excellent	Average Measures	0.910[c]	0.858	0.943	11.406	74	74	<0.001
Distal	ExoCad	TM	1st	Excellent	Single Measures	0.867[a]	0.797	0.914	13.866	74	74	<0.001
		TA	1st	Excellent	Average Measures	0.929[c]	0.887	0.955	13.866	74	74	<0.001
Buccal	ExoCad	TM	1st	Good	Single Measures	0.685[a]	0.539	0.790	5.622	74	74	<0.001
		TA	1st	Excellent	Average Measures	0.813[c]	0.700	0.883	5.622	74	74	<0.001
Lingual	ExoCad	TM	1st	Good	Single Measures	0.731[a]	0.606	0.821	6.477	74	74	<0.001
		TA	1st	Excellent	Average Measures	0.845[c]	0.755	0.902	6.477	74	74	<0.001
Mesial	Fusion 360	TA	2nd	Poor	Single Measures	0.207[a]	−0.005	0.405	1.570	74	74	0.027
		TM	2nd	Poor	Average Measures	0.343[c]	−0.011	0.577	1.570	74	74	0.027
Distal	Fusion 360	TA	2nd	Fair	Single Measures	0.405[a]	0.185	0.584	2.593	74	74	<0.001
		TM	2nd	Fair	Average Measures	0.577[c]	0.313	0.737	2.593	74	74	<0.001
Buccal	Fusion 360	TA	2nd	Poor	Single Measures	0.387[a]	0.176	0.564	2.252	74	74	<0.001
		TM	2nd	Fair	Average Measures	0.558[c]	0.299	0.721	2.252	74	74	<0.001
Lingual	Fusion 360	TA	2nd	Poor	Single Measures	0.166[a]	−0.051	0.372	1.418	74	74	0.068
		TM	2nd	Poor	Average Measures	0.285[c]	−0.108	0.542	1.418	74	74	0.068
Mesial	ExoCad	TA	2nd	Poor	Single Measures	0.258[a]	0.033	0.458	1.688	74	74	0.013
		TM	2nd	Fair	Average Measures	0.410[c]	0.064	0.628	1.688	74	74	0.013
Distal	ExoCad	TA	2nd	Poor	Single Measures	0.296[a]	0.083	0.486	1.956	74	74	0.002
		TM	2nd	Fair	Average Measures	0.457[c]	0.153	0.654	1.956	74	74	0.002
Buccal	ExoCad	TA	2nd	Fair	Single Measures	0.412[a]	0.204	0.584	2.385	74	74	<0.001
		TM	2nd	Fair	Average Measures	0.583[c]	0.339	0.737	2.385	74	74	<0.001
Lingual	ExoCad	TA	2nd	Fair	Single Measures	0.429[a]	0.226	0.596	2.623	74	74	<0.001
		TM	2nd	Good	Average Measures	0.600[c]	0.369	0.747	2.623	74	74	<0.001

[a]The estimator is the same, whether the interaction effect is present or not. [b]Type A intraclass correlation coefficients using an absolute agreement definition. [c]This estimate is computed assuming the interaction effect is absent, because it is not estimable otherwise.

Table 5: Interclass correlation coefficient showing Inter- rater reliability within the software and between the 1st time and 2nd time of measurement

Undercut type/ location	Software (s)	Rater (s)	Rating time (1st or 2nd)	Interpretation of intraclass correlation coefficient	Reliability	Intraclass correlation[b]	95% confidence interval		F-test with true value-0			
							Lower bound	Upper bound	Value	df1	df2	Sig
Mesial	Fusion 360	TM	1st	Poor	Single Measures	0.255[a]	0.040	0.450	1.819	74	74	0.005
		TA	2nd	Fair	Average Measures	0.406[c]	0.076	0.620	1.819	74	74	0.005
Distal	Fusion 360	TM	1st	Poor	Single Measures	0.296[a]	0.083	0.486	1.956	74	74	0.002
		TA	2nd	Fair	Average Measures	0.457[c]	0.153	0.654	1.956	74	74	0.002
Buccal	Fusion 360	TM	1st	Poor	Single Measures	0.302[a]	0.090	0.490	1.931	74	74	0.003
		TA	2nd	Fair	Average Measures	0.464[c]	0.165	0.658	1.931	74	74	0.003
Lingual	Fusion 360	TM	1st	Poor	Single Measures	0.222[a]	0.006	0.421	1.741	74	74	0.009
		TA	2nd	Poor	Average Measures	0.363[c]	0.013	0.593	1.741	74	74	0.009
Mesial	ExoCad	TM	1st	Poor	Single Measures	0.369[a]	0.155	0.549	2.159	74	74	<0.001
		TA	2nd	Fair	Average Measures	0.539[c]	0.269	0.709	2.159	74	74	<0.001
Distal	ExoCad	TM	1st	Poor	Single Measures	0.278[a]	0.066	0.469	1.844	74	74	0.005
		TA	2nd	Fair	Average Measures	0.435[c]	0.123	0.638	1.844	74	74	0.005
Buccal	ExoCad	TM	1st	Poor	Single Measures	0.297[a]	0.084	0.487	1.889	74	74	0.003
		TA	2nd	Fair	Average Measures	0.459[c]	0.155	0.655	1.889	74	74	0.003
Lingual	ExoCad	TM	1st	Poor	Single Measures	0.376[a]	0.169	0.553	2.295	74	74	<0.001
		TA	2nd	Fair	Average Measures	0.547[c]	0.289	0.712	2.295	74	74	<0.001
Mesial	Fusion 360	TA	1st	Poor	Single Measures	0.255[a]	0.040	0.450	1.819	74	74	0.005
		TM	2nd	Fair	Average Measures	0.406[c]	0.076	0.620	1.819	74	74	0.005
Distal	Fusion 360	TA	1st	Poor	Single Measures	0.296[a]	0.083	0.486	1.956	74	74	0.002
		TM	2nd	Fair	Average Measures	0.457[c]	0.153	0.654	1.956	74	74	0.002
Buccal	Fusion 360	TA	1st	Poor	Single Measures	0.302[a]	0.090	0.490	1.931	74	74	0.003
		TM	2nd	Fair	Average Measures	0.464[c]	0.165	0.658	1.931	74	74	0.003
Lingual	Fusion 360	TA	1st	Poor	Single Measures	0.222[a]	0.006	0.421	1.741	74	74	0.009
		TM	2nd	Poor	Average Measures	0.363[c]	0.013	0.593	1.741	74	74	0.009
Mesial	ExoCad	TA	1st	Good	Single Measures	0.644[a]	0.489	0.759	4.573	74	74	<0.001
		TM	2nd	Excellent	Average Measures	0.783[c]	0.657	0.863	4.573	74	74	<0.001
Distal	ExoCad	TA	1st	Good	Single Measures	0.644[a]	0.489	0.759	4.573	74	74	<0.001
		TM	2nd	Excellent	Average Measures	0.783[c]	0.657	0.863	4.573	74	74	<0.001

(*Contd*...)

Table 5: (*Continued*)

Undercut type/location	Software (s)	Rater (s)	Rating time (1^{st} or 2^{nd})	Interpretation of intraclass correlation coefficient	Reliability	Intraclass correlation^{b}	Intraclass correlation coefficient: 95% confidence interval Lower bound	Upper bound	F-test with true value-0: Value	df1	df2	Sig
Buccal	ExoCad	TA	1^{st}	Good	Single Measures	0.644^{a}	0.489	0.759	4.573	74	74	<0.001
		TM	2^{nd}	Excellent	Average Measures	0.783^{c}	0.657	0.863	4.573	74	74	<0.001
Lingual	ExoCad	TA	1^{st}	Good	Single Measures	0.644^{a}	0.489	0.759	4.573	74	74	<0.001
		TM	2^{nd}	Excellent	Average Measures	0.783^{c}	0.657	0.863	4.573	74	74	<0.001

Two-way mixed effects model where people effects are random and measures effects are fixed. [a]The estimator is the same, whether the interaction effect is present or not. [b]Type A intraclass correlation coefficients using an absolute agreement definition. [c]This estimate is computed assuming the interaction effect is absent, because it is not estimable otherwise

undercut's presence between the 1^{st}-time and the 2^{nd}-time rating for each software. Rater TM had excellent ICC reliability of the presence of undercut between his 1^{st} and 2^{nd} time in Exocad and Fusion 360. Rater TA had poor to fair ICC reliability of the presence of undercut between his 1^{st} and 2^{nd} time in Exocad and Fusion 360, Table 1.

When checking the intra-rater reliability details of each rater were compared in their reliability of rating of undercuts per location between the 1^{st}-time and 2^{nd}-time rating, Rater TM has poor to fair reliability between the 1^{st} and 2^{nd} time in almost all the locations and software, except in distal locations using Exocad. While rater TA had good to excellent reliability between the 1^{st} and 2^{nd} time in almost all the locations of Fusion 360 but fair to poor reliability in all locations using Exocad, Table 2.

For inter-rater reliability, both raters, TM and TA, were compared to each other for the presence of undercuts on the reliability in all situation's combinations of software, raters, and rating instances (1^{st} or 2^{nd} time) shown in Table 3. Utilizing each software during the same instance, first to first, the two raters, TM and TA, had excellent reliability or agreement. However, during the second round of rating, TM and TA had excellent agreement only with Exocad. In the mixed rating instances, excellent agreements were found in the last two combinations in Table 3.

If looking within each software separately by location, Inter- rater reliability within the software and between the 1^{st} times or 2^{nd} times of measurement Table 4. TM and TA had excellent agreement and reliability using Exocad during the 1^{st} time in all locations. Other than that, the agreement was poor to fair.

If looking within each software separately by location, inter- rater reliability within the software and between the 1^{st} times and 2^{nd} times of measurement Table 5. TM and TA had excellent agreement and reliability only when using Exocad during the 1^{st} time TA and 2^{nd} time TM in all locations. Other than that, the agreement was poor to fair.

Undercut counts by raters and software are shown in Table 6 and Figure 1. While undercut counts and percentages by raters, locations and software were shown in Table 7; Figures 2 and 3.

Detailed ICC, 95% confidence intervals, significance, and levels of inter-rater reliability are shown in Tables 1-5.

The current study shows the presence of undercuts ranging from 85.3% to 98.7% when measured by both software and the two raters [Table 6]. Detailed counts and percentages of undercuts for each factor are shown in Tables 6 and 7.

Table 6: Undercut counts and percentages by raters and software

Operator	Fusion Count	Fusion Column *n* %	Exocad Count	Exocad Column *n* %
Rater TM1 Undercut	65	86.70	64	85.30
Rater TM1 No Undercut	10	13.30	11	14.70
Rater TA1 Undercut	70	93.30	68	90.70
Rater TA1 No Undercut	5	6.70	7	9.30
Rater TM2 Undercut	64	85.30	63	84.00
Rater TM2 No Undercut	11	14.70	12	16.00
Rater TA2 Undercut	74	98.70	74	98.70
Rater TA2 No Undercut	1	1.30	1	1.30

Table 7: Undercut counts and percentages by raters, locations and software

Operator	Fusion Count	Fusion Column *n* %	Exocad Count	Exocad Column *n* %
TM1 Mesial Undercut	26	34.70	43	57.30
TM1 Distal Undercut	39	52.00	41	54.70
TM1 Buccal Undercut	29	38.70	31	41.30
TM1 Lingual Undercut	22	29.30	34	45.30
TA1 Mesial Undercut	46	61.30	47	62.70
TA1 Distal Undercut	50	66.70	42	56.00
TA1 Buccal Undercut	46	61.30	39	52.00
TA1 Lingual Undercut	42	56.00	30	40.00
TM2 Mesial Undercut	31	41.30	43	57.30
TM2 Distal Undercut	39	52.00	39	52.00
TM2 Buccal Undercut	39	52.00	40	53.30
TM2 Lingual Undercut	34	45.30	34	45.30
TA2 Mesial Undercut	46	61.30	46	61.30
TA2 Distal Undercut	55	73.30	55	73.30
TA2 Buccal Undercut	42	56.00	42	56.00
TA2 Lingual Undercut	46	61.30	46	61.30

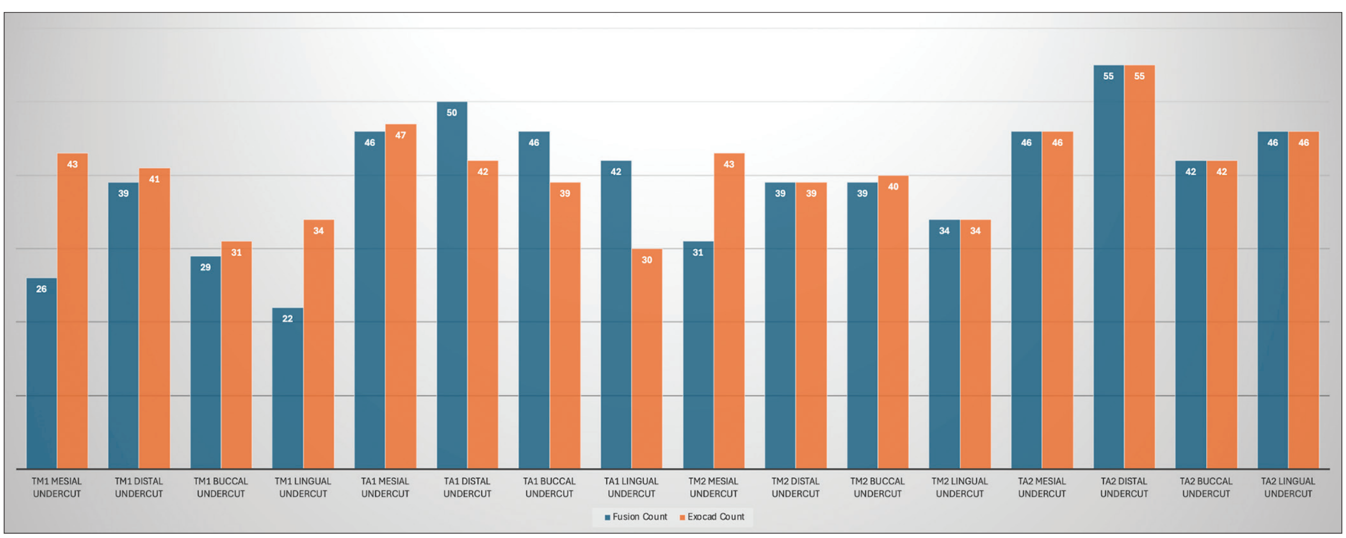

Figure 2: Undercut counts by raters, location, and software

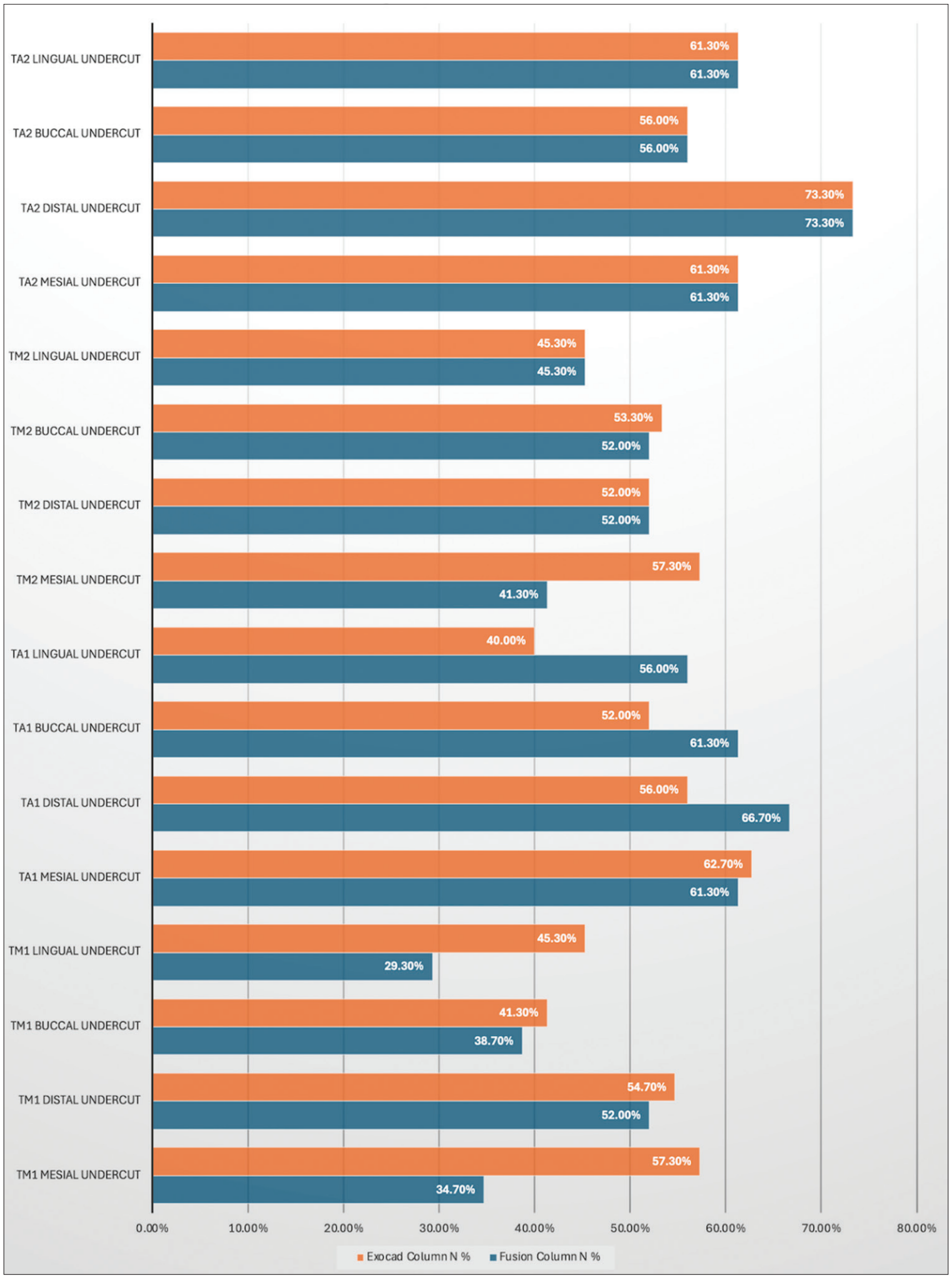

Figure 3: Undercut percentage by raters, location, and software

DISCUSSION

When evaluating the reliability of the two software in detecting the undercut's specific location by each separate rater, both software had poor reliability (agreement).

That could be attributed to the ease of use in selecting the POI in the ExoCad software compared to Fusion 360. Therefore, each rater could not specify the location of the undercut when using the software. This study utilized single crown preparations. Four walls are involved in determining the path of crown insertion. All four axial walls collectively dictate the path of crown insertion, generating a broader spectrum of permissible paths of insertion. This expanded range inherently increases the variability in undercut presentation, thereby reducing consistency in identifying their exact location. This explains the poor agreement in determining the specific location of the undercut. Undercuts are far more important when there are multiple teeth involved. The more vertical walls and planes involved in the preparations that share a common POI, the more specific the POI is and the less the range of variation in selecting the common POI between the operators. There is a need to find a unified approach to detecting the POI of teeth using CAD software algorithms to detect undercuts in dental teeth preparations properly. That need becomes more pronounced when multiple teeth and surfaces are involved in teeth preparations where undercut elimination is crucial for proper seating and retention of the final restorations.

CONCLUSION

Variations between the software in the detection of undercut are operator related. The ease of determining the proper alignment of the POI done by the operator before measurement is crucial for the agreement and reliability of the detection of undercuts.

FUTURE DIRECTION

Future studies involve the evaluation of undercuts in multiunit FDP preparations.

ETHICAL APPROVAL

The research proposal was exempted by the Institutional Review Board of King Abdulaziz University, as the study was conducted as an *in vitro* experiment.

ACKNOWLEDGMENT

Special thanks to the Deanship of Scientific Research and the faculty of Dentistry at King Abdulaziz University, Jeddah, Saudi Arabia, for supporting this project.

DATA AND MATERIALS AVAILABILITY

All data associated with this study are present in the paper. *This study followed EQUATOR guidelines.

REFERENCES

1. Hey J, Schweyen R, Kupfer P, Beuer F. Influence of preparation design on the quality of tooth preparation in preclinical dental education. J Dent Sci 2017;12:27-32.
2. Marghalani TY. Frequency of undercuts and favorable path of insertion in abutments prepared for fixed dental prostheses by preclinical dental students. J Prosthet Dent 2016;116:564-9.
3. Ayad MF, Maghrabi AA, Rosenstiel SF. Assessment of convergence angles of tooth preparations for complete crowns among dental students. J Dent 2005;33:633-8.
4. El-Mowafy OM, Fenton AH, Forrester N, Milenkovic M. Retention of metal ceramic crowns cemented with resin cements: Effects of preparation taper and height. J Prosthet Dent 1996;76:524-9.
5. Cicchetti DV. Multiple comparison methods: Establishing guidelines for their valid application in neuropsychological research. J Clin Exp Neuropsychol 1994;16:155-61.
6. McGraw KO, Wong SP. Forming inferences about some intraclass correlation coefficients. Psychol Methods 1996;1:30-46.
7. Shrout PE, Fleiss JL. Intraclass correlations: Uses in assessing rater reliability. Psychol Bull 1979;86:420-8.

Source of Support: This project was funded by the Deanship of Scientific Research (DSR) at King Abdulaziz University (KAU), Jeddah, under grant no. J:016-165-1443. **Conflicts of Interest:** None declared.