

The Impact of Varying Levels of AI Feedback on Novice Clinicians' Confidence and Accuracy in Detecting Furcation Involvement

Ariana Bagheri, Jennifer Chang, DDS, MSD, Aaron Glick, DDS

Objectives

Recent advancements in artificial intelligence offer promising enhancements to diagnostic accuracy in dentistry. AI tools are being explored for automating the detection of dental conditions like furcation involvement. This study aims to determine the effect of varying levels of perceived AI accuracy feedback on students' performance and confidence in detecting furcation involvement on radiographic images.

Experimental methods:

Eighty students were randomly assigned to one of 4 groups with varying AI feedback: 60%, 80%, 99%, and a control group (with no AI). Each group completed a 17-question survey via the Qualtrics platform (Table 1).

Statistical analyses: ANOVA tests were performed for time, accuracy, and agreement with AI between groups. Paired t-tests were performed for both self-confidence and confidence in AI pre- and post-survey. A 5% significance level was used for all tests.

Results:

Significant differences on the total average time spent on furcation involvement detection were found between groups ($p < 0.001$), with higher AI accuracy feedback leading to less time spent per question, except for the control group. The 99% group had a total time of 1943.62 ± 28.01 seconds spent on the survey, the 80% group had a total time of 2438.19 ± 58.83 , the 60% group had a total time 2776.61 ± 69.10 , and the control group had a total time of 1477.32 ± 46.48 time spent on the survey. (Table 2). No significant differences were found in accuracy or confidence levels across the groups (Table 3-5).

Conclusion:

Perceptions of more accurate AI systems were associated with shorter average time spent identifying furcation involvement. However, perceived AI accuracy did not significantly affect students' detection accuracy or confidence, suggesting that perceived accuracy alone may not build clinician-AI trust.

This study was supported by the UTSD Student Research Program.

Table 1.

Qualtrics Survey





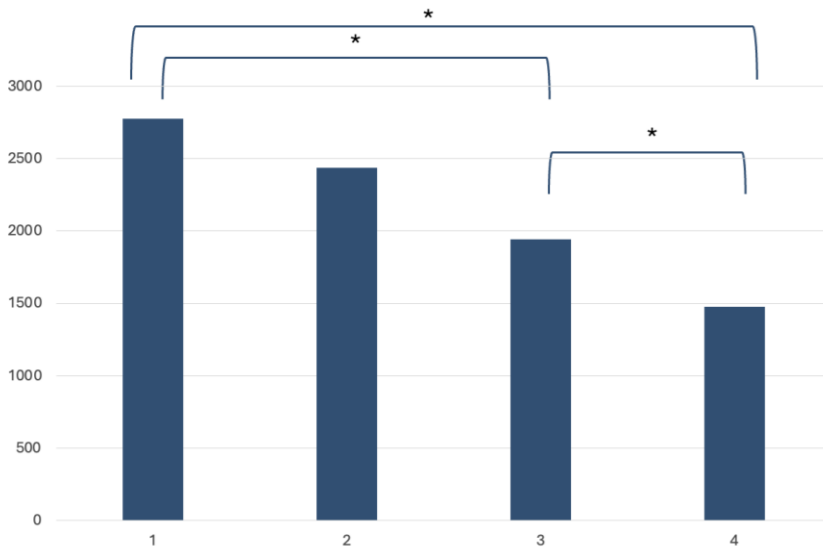
Self			
What's your year of training? Before we get started, what's your confidence level on radiographic identification of furcation lesions?			
AI			
Do you think practicing with AI will help your clinical confidence?			
Furcation involvement (Example)			
60%	80%	99%	Control (No AI)
			
AI identified furcation lesion on tooth #2 (Accuracy of AI was 60%) Please identify if there is a furcation lesion on tooth #2 (molar on the left of the image)	AI identified furcation lesion on tooth #2 (Accuracy of AI was 80%) Please identify if there is a furcation lesion on tooth #2 (molar on the left of the image)	AI identified furcation lesion on tooth #2 (Accuracy of AI was 99%) Please identify if there is a furcation lesion on tooth #2 (molar on the left of the image)	Please identify if there is a furcation lesion on tooth #2 (molar on the left of the image)
Post			
Self			
What's your confidence level on your diagnostic accuracy of identifying previous radiographic furcation lesions?			
AI			
Do you think a computer program for practicing on radiographic furcation lesion identification would help your clinical confidence?			

Table 2.

Total Time Spent on Furcation Identification Questions

60%	80%	99%	Control
2776.61±69.10	2438.19±58.83	1943.62±28.01	1477.32±46.48



A significant difference was found among all (p-value=<0.01)

*Significant difference found between found between groups (p-value=<0.05)

Table 3.

Furcation Involvement Identification Accuracy, Agreement with AI, Time Spent

AI Answer	Yes				No				Yes				No				Yes				Yes				No				No				No				No											
Gold Standard	Yes				No				Yes				No				Yes				Yes				No				No				No				No											
Agreement with AI	60	80	99	C	60	80	99	C	60	80	99	C	60	80	99	C	60	80	99	C	60	80	99	C	60	80	99	C	60	80	99	C	60	80	99	C	60	80	99	C	60	80	99	C	60	80	99	C
Accuracy	5%	10%	15%	10%	55%	45%	55%	45%	0%	5%	5%	5%	10%	25%	10%	10%	45%	55%	55%	70%	25%	12%	30%	10%	0%	0%	0%	0%	5%	5%	0%	10%	5%	5%	0%	10%	5%	5%	0%	10%	0%	15%	0%	10%	0%	0%	0%	0%
Time	5%	10%	15%	10%	45%	55%	45%	55%	0%	5%	5%	5%	10%	25%	10%	10%	45%	55%	55%	70%	25%	12%	30%	10%	0%	0%	0%	0%	5%	5%	0%	10%	5%	5%	0%	10%	0%	15%	0%	10%	0%	0%	0%	0%				
	13.46	12.84	10.56	5.76	20.83	15.94	13.30	13.28	10.16	8.8	9.12	6.84	12.44	8.54	7.09	4.5	23.3	19.61	12.33	8.83	9	8.33	6.1	4.04	16.82	18.45	14.53	12.49	14.56	9.82	7.77	7.24	10.2	13.82	10.7	6.72	8.05	5.75	5.68	4.12								

Groups: 60 (AI with 60% of accuracy), 80 (AI with 80% of accuracy), 99 (AI with 99% of accuracy), C (Control: No AI given)

Table 4.

Results on Self Confidence Before and After Survey

60		80		99		Control	
Pre	Post	Pre	Post	Pre	Post	Pre	Post
3.35±0.93	2.90 ±0.79	3.45±0.76	3.60±0.75	3.45± 0.69	3.15±0.81	3.60±0.75	3.25 ±0.55
P=0.06		P=0.33		P=0.16		P=0.11	

No significant difference among all groups on self confidence level (ranged 1-5, 1 indicated minimum confidence, 5 indicated maximum confidence) before and after survey completion

Table 5.

Results on Confidence in AI Before and After Survey

60		80		99		Control	
Pre	Post	Pre	Post	Pre	Post	Pre	Post
3.70±0.86	3.95 ±0.83	3.95±0.83	4.10±0.45	4.40 ± 0.75	4.20±0.83	4.10±0.55	4.00 ±0.92
P=0.10		P=0.53		P=0.30		P=0.54	

No significant difference among all groups on confidence level in AI (ranged 1-5, 1 indicated minimum confidence, 5 indicated maximum confidence) before and after survey completion