

Validation Report

Clinical validation of Organic Plaque Simulation (PG Plaque) in Robot Toothbrushing Tests

Witten

2020



ORMED Institute for Oral Medicine at the University
of Witten/ Herdecke

Prof. Dr. Dr. h.c. P. Gängler
Dr. Tomas Lang
Dr. Karl Weich
Henrike Pepin
Christian Greune

Content

1	Abstract	3
2	Methods and Test Materials	4
3	Results	7
	<i>3.1 Comparisons between toothbrushes</i>	<i>7</i>
	<i>3.2 Assessment of Agreement Rate AR</i>	<i>10</i>
4	Statistical analysis	14
	<i>4.1 Kolmogorov-Smirnov-Test</i>	<i>14</i>
	<i>4.2 Independent t-Test</i>	<i>15</i>
	<i>4.3 Wilcoxon-Mann-Whitney-U-Test</i>	<i>20</i>
5	Conclusions	22
6	References	23
7	Appendix	24
	<i>7.1 Single values of clinical testing</i>	<i>24</i>
	<i>7.2 Single values of robot testing</i>	<i>25</i>
	<i>7.3 Single values tooth by tooth comparing clinical and robot testing</i>	<i>26</i>

1 Abstract

Objectives

Robot testing of simulated plaque control is important for developing new toothbrushes and full mouth devices. Therefore, the aim was (i) to test a novel formulation of organic plaque simulating viscosity and adhesion of natural plaque and (ii) to estimate a valid accuracy of robot outcome in relation to clinical results of plaque control.

Methods

Clinical programme: After ethical approval (EK-UWH 552007), professional tooth cleaning and 3-day-plaque-regrowth was executed, and 22 calibrated subjects used in a Randomized Clinical Trial video-supported separated horizontal, and rotating, and vertical brushing movements for 20s buccally/20s lingually at 9 teeth 32 – 47 with force 3.5 N. Toothbrushes Dr. Best medium (TB1) and Interdent medium (TB2) (GlaxoSmithKline, Munich, Germany) were tested. Stained plaque was photographed and blind-coded at 18 planimetric fields and at 10 risk fields using modified Navy-Plaque-Index (Lang et al. 2011) with PPI-Codes 0 (0%), 1(<50 %) and 2(>50 %) per each field.

Robot programme: The same brushes and techniques were tested. The cleaning outcome of simulated organic plaque in percentage per planimetric field with Computer-assisted Planimetric Plaque Assessment (APP) was blind-assessed with PPI.

All clinical and robot data underwent statistical analysis by K-S-test, one-sample-t-test, Independent t-test of equality of means, W-M-W-U-Test of equality of medians and Agreement Rate AR of plaque removal.

Results

Individual clinical plaque control pattern at two surfaces and two risk areas per tooth were well reproduced by robot brushing movements. The Agreement Rate of plaque removal by separated brushing movements at smooth surfaces was 85–100 % (TB1) and 89–99 % (TB2); at risk fields next to gum line 84–98 % (TB1) and 88–94 % (TB2). The single tooth analysis revealed best AR for teeth 42 (TB1 83-99%), 42 (TB2 81-98%) and 47 (TB2 75-98%). Canines 43 exhibited the least AR 41% for both brushes. All 24 tooth sites (buccally and lingually) and all risk areas exhibited in all 3 brushing movements with the 2 toothbrushes equal plaque control values ($p = 0.05$) or, alternatively, 21 out of 24 tooth sites showed equal values (most common $p = 0.10$).

Conclusions

Robot toothbrushing with the formulation of organic plaque, simulating bio-physical parameters of natural plaque, is concordant with clinical plaque control at all teeth and all planimetric areas. The clinically validated plaque simulation is recommended for complex dry and wet robot testing.

2 Methods and Test Materials



Fig. 1-4: Stained clinical plaque after 3-day plaque regrowth (Fig. 1); stained clinical plaque after brushing for 20 s with force 3.5 N (Fig.2); stained organic plaque simulation on KaVo teeth (Fig. 3); stained organic plaque simulation after robot brushing for 20 s with force 3.5 N (Fig. 4)

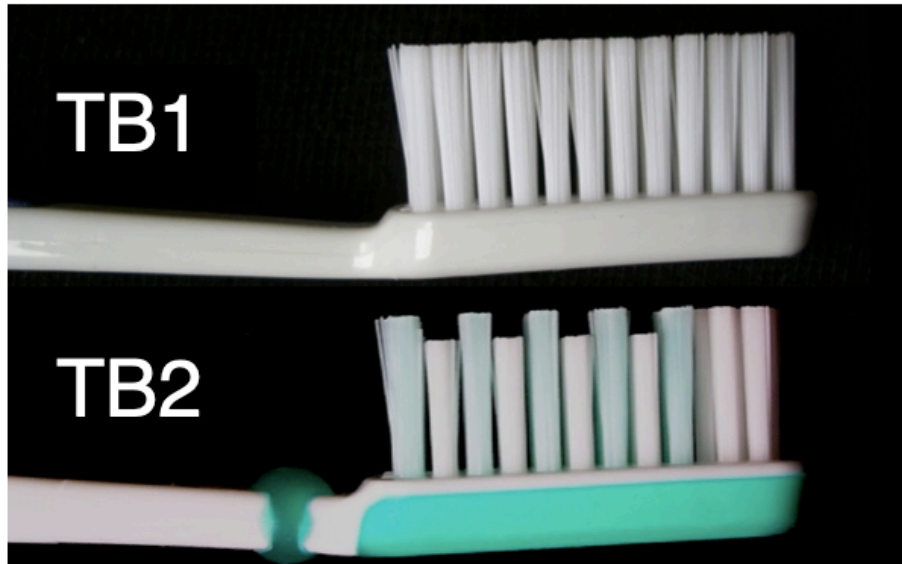


Fig. 5: Toothbrushes Dr. Best Flat Cut medium (TB 1) and Interdent medium (TB 2)



Fig. 6: Automated Plaque Planimetry (APP); teeth covered with organic plaque simulation after brushing, site by site rotating in front of the HD focusing analysis camera followed by computer-assisted processing of plaque percentage per each single field

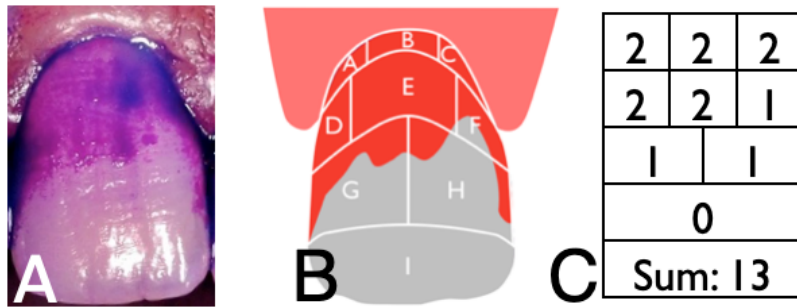


Fig. 7: Planimetric fields at human teeth (A), clinical brushing outcome (B), Planimetric Plaque Index PPI Scores (Lang et al. 2011) (C)

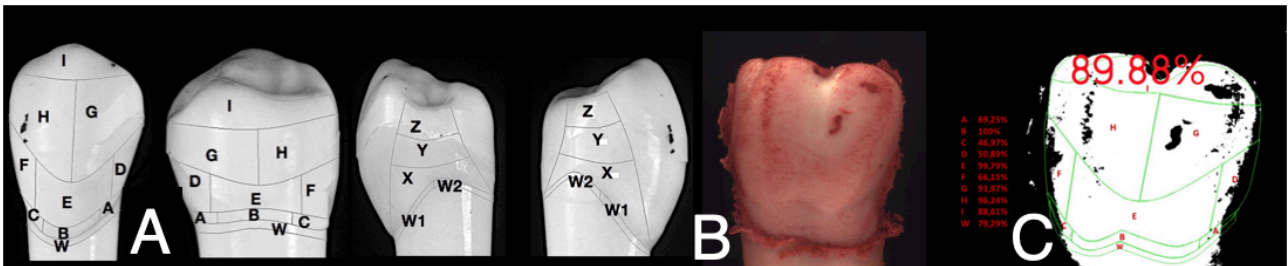


Figure 8: Planimetric fields at human teeth (A), in-vitro brushing outcome on molar tooth (B), Automated Planimetric Plaque Assessment (APP) (C)

3 Results

3.1 Comparisons between toothbrushes

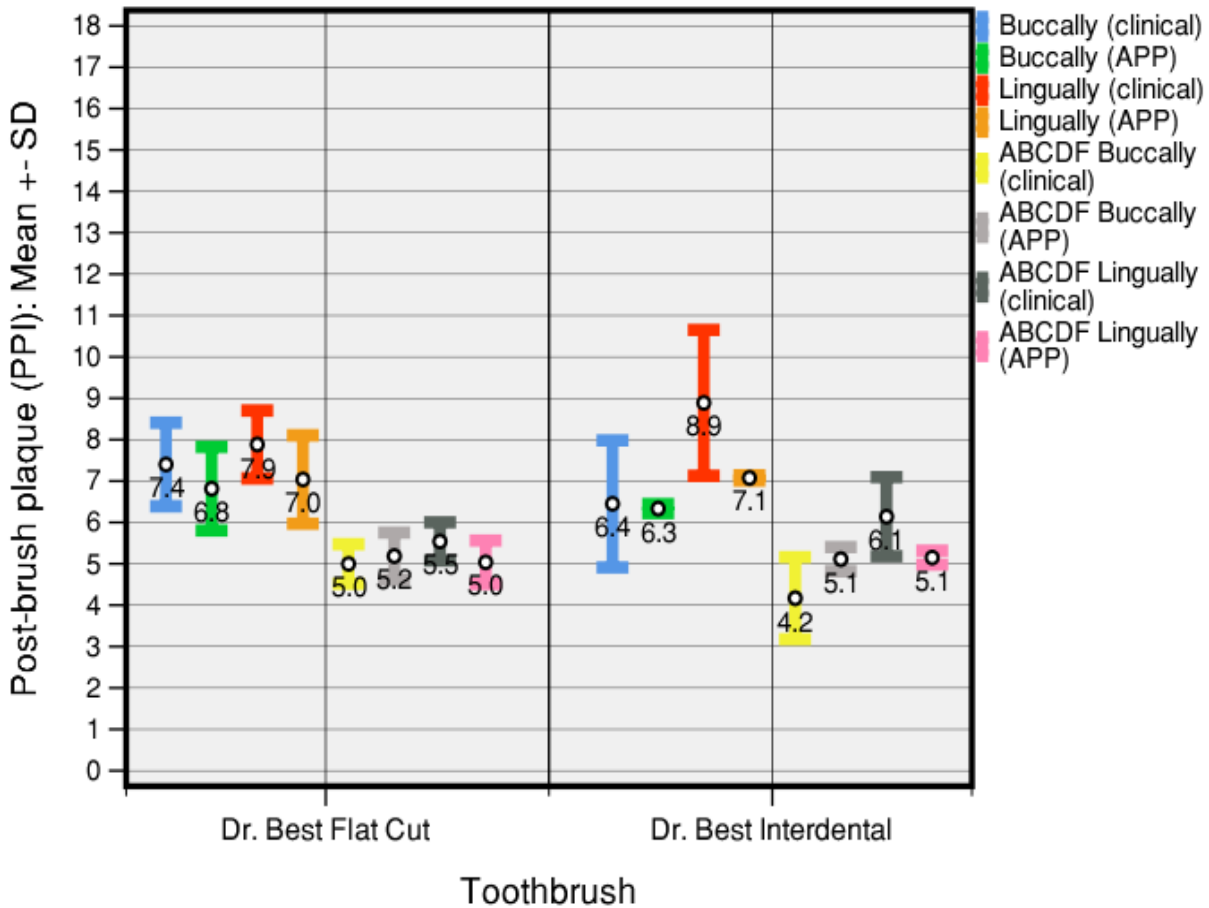


Fig. 9: horizontal brushing movement

Error bars of post-brush plaque (Modified Navy Plaque Index according to Lang et al. 2011 at clinical values PPI and at transformed APP values) buccally smooth surfaces (towards the cheek), lingually smooth surfaces (towards the tongue) and at buccal and lingual risk fields ABCDF (next to the gum line) for the two tested toothbrushes

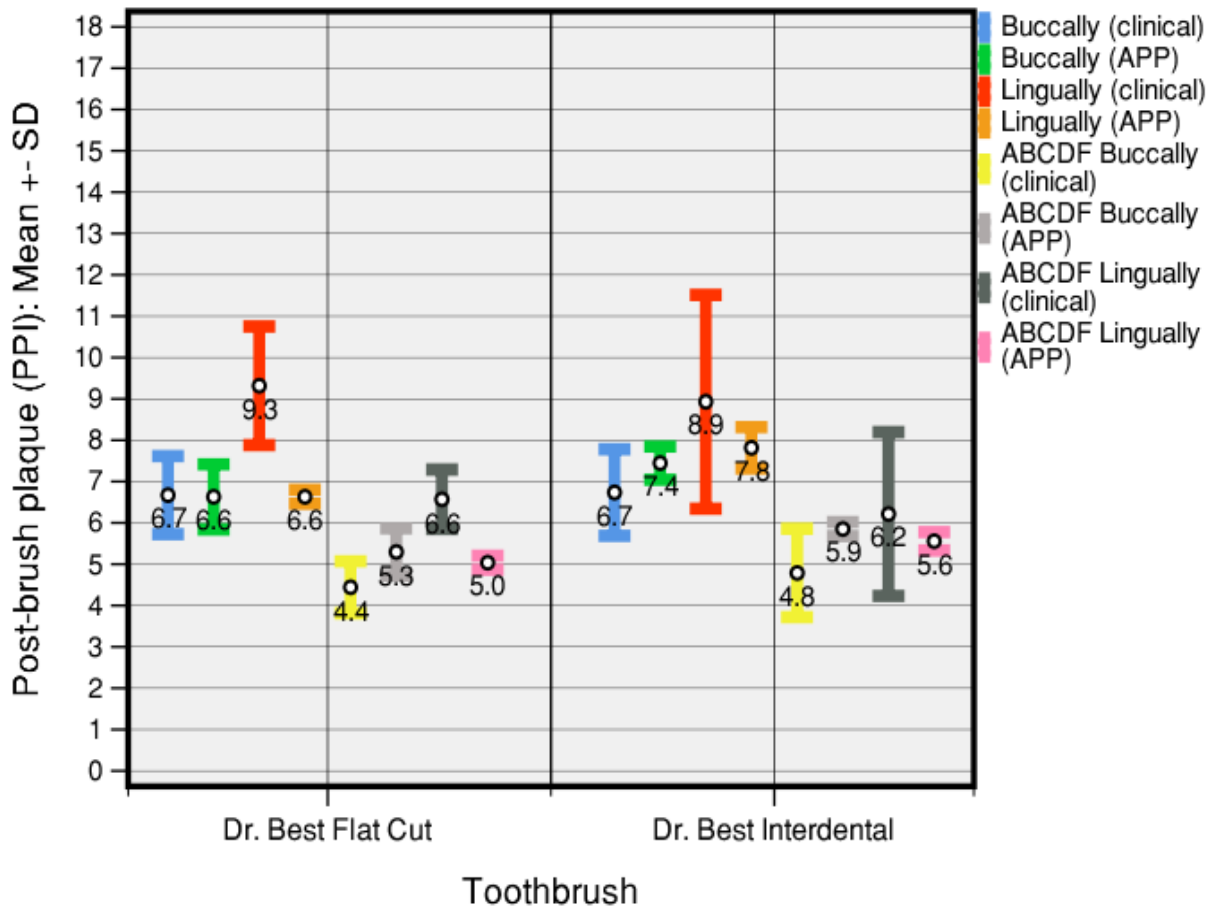


Fig. 10: rotating brushing movement

Error bars of post-brush plaque (Modified Navy Plaque Index according to Lang et al. 2011 at clinical values PPI and at transformed APP values) buccally smooth surfaces (towards the cheek), lingually smooth surfaces (towards the tongue) and at buccal and lingual risk fields ABCDF (next to the gum line) for the two tested toothbrushes

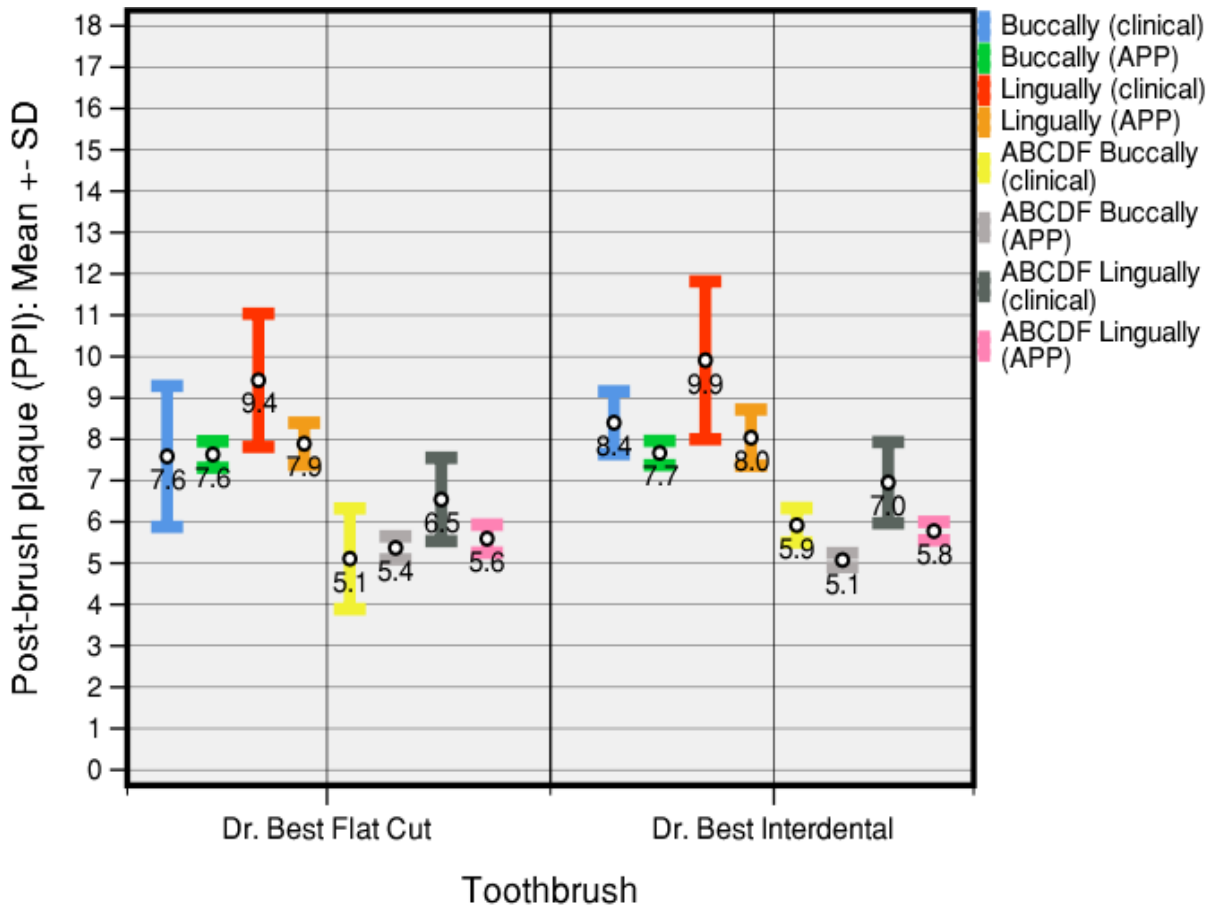


Fig. 11: vertical brushing movement

Error bars of post-brush plaque (Modified Navy Plaque Index according to Lang et al. 2011 at clinical values PPI and at transformed APP values) buccally smooth surfaces (towards the cheek), lingually smooth surfaces (towards the tongue) and at buccal and lingual risk fields ABCDF (next to the gum line) for the two tested toothbrushes

3.2 Assessment of Agreement Rate AR

The buccal and lingual sites of each tooth are divided in 9 planimetric fields A - I according to the Navy Plaque Index and coded 0 = no plaque, 1 = plaque coating less than 50 % of the field, 2 = plaque coating 50 and more % of the field (PPI , Lang et al. 2011).

The max. code value per site is, therefore, 18 and all 9 fields are coded 2.

The max. difference per tooth site comparing clinical results with the robot test data APP is, consequently, 18 and the Agreement Rate AR = 0 %.

No difference per tooth site comparing clinical results with the robot test data is, consequently, the Agreement Rate AR = 100 %.

Comparing the risk areas ABCDF next to the gum line, the Agreement Rate AR is counted differently. The max. difference is 10, therefore, the Agreement Rate is AR = 0 %. In case of no difference, the AR is again 100 %.

Examples:

1.	PPI clin (vertically, buccally, Flat Cut)	=	7.6	
	PPI APP (vertically, buccally, Flat Cut)	=	7.6	
	Δ (PPI clin vs. PPI APP)	=	$7.6 - 7.6$	= 0
	AR (PPI clin vs. PPI APP)	=	$(18-0) / 18 = 1$	= 100%
2.	PPI clin (horizontally, lingually, Dr. Best Interdent)	=	8.9	
	PPI APP (horizontally, lingually, Dr. Best Interdent)	=	7.1	
	Δ (PPI clin vs. PPI APP)	=	$8.9 - 7.1 = 1.8$	
	AR (PPI clin vs. PPI APP)	=	$(18-1.8) / 18 = 0.9$	= 90%
3.	PPI clin (ABCDF, rotating, buccally, Flat Cut)	=	4.4	
	PPI APP (ABCDF, rotating, buccally, Flat Cut)	=	5.8	
	Δ (PPI clin vs. PPI APP)	=	$4.4 - 5.8$	= 1.4
	AR (PPI clin vs. PPI APP)	=	$(10-1.4) / 10 = 0.86$	= 86%
4.	PPI clin (Tooth 31, buccally, horizontally, Flat Cut)	=	6.13	
	PPI APP (Tooth 31, buccally, horizontally, Flat Cut)	=	6	
	Δ (PPI clin vs. PPI APP)	=	$6.13 - 6$	= 0.13
	AR (PPI clin vs. PPI APP)	=	$(18-0.13) / 18 = 0.993$	= 99.3%
5.	PPI clin (Tooth 43, ABCDF ling., horiz., Flat Cut)	=	3.49	
	PPI APP (Zahn 43, ABCDF ling., horiz., Flat Cut)	=	9.3	
	Δ (PPI clin vs. PPI APP)	=	5.84	
	AR (PPI clin vs. PPI APP)	=	$(10-5.84) / 10 = 0.416$	= 41.6%

Total planimetric fields, risk areas, single tooth analyses

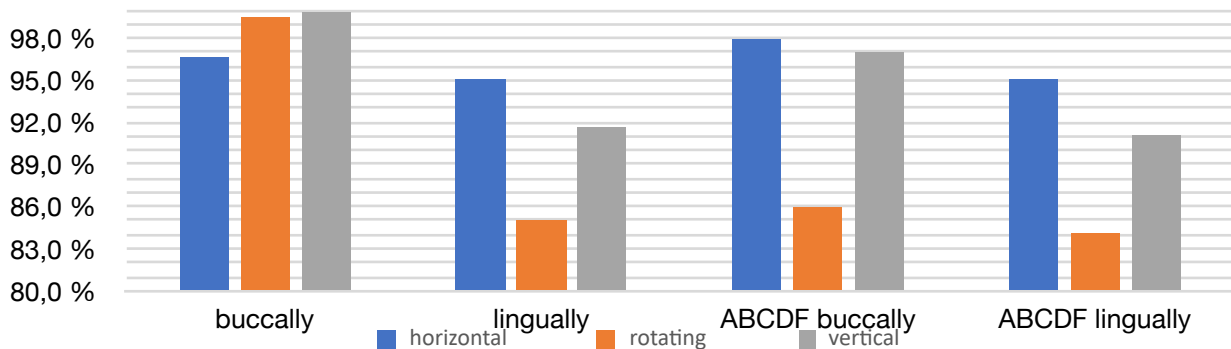
Agreement Rate Dr. Best Flat Cut
clin with Dr. Best Flat Cut APP in %
(($\Delta = 0$) = (AR = 100%))

	Horizontal	Rotating	Vertical
buccally	96.7%	99.4%	100.0%
lingually	95.0%	85.0%	91.7%
ABCDF buccally (risk field)	98.0%	86.0%	97.0%
ABCDF lingually (risk field)	95.0%	84.0%	91.0%

Agreement Rate Dr. Best Interdent
clin with Dr. Best Interdent APP in %
(($\Delta = 0$) = (AR = 100%))

	Horizontal	Rotating	Vertical
buccally	99.4%	96.1%	96.1%
lingually	90.0%	93.9%	89.4%
ABCDF buccally (risk field)	91%	89%	92%
ABCDF lingually (risk field)	90%	94%	88%

Agreement Rate Dr. Best Flat Cut clin with Dr. Best Flat Cut APP in %
Delta (0)= AR(100%) in view of Fig 9-11



Agreement Rate Dr. Best Interdent clin with Dr. Best Interdent APP in %
Delta (0)= AR(100%) in view of Fig 9-11

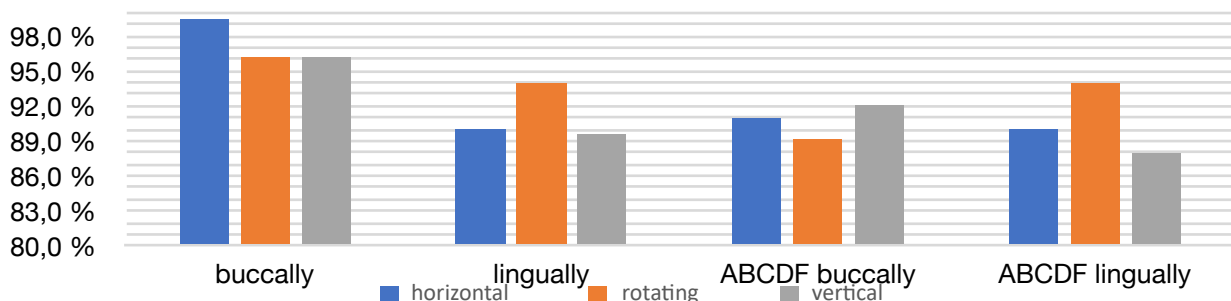


Fig. 12: The Agreement Rate is highest at buccal sites for both brushes, lowest at lingual sites for both brushes; horizontal brushing shows the best AR, followed by vertical brushing, and, last but not least by rotating brushing.

Agreement Rate Dr. Best Flat Cut clin with Dr. Best Flat Cut APP in % Delta (0) = AR (100%) in view of Tab. 5a, 5b, 5c

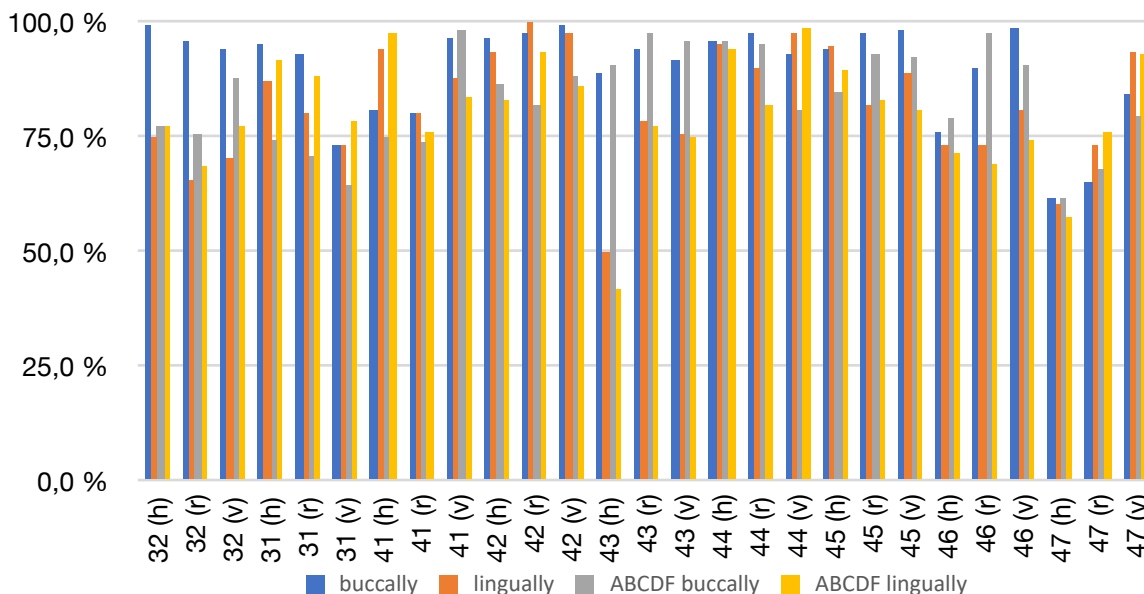


Fig. 13: Agreement Rate AR of toothbrushing with TB 1; AR per single incisors, canines, premolars and molars (h – horizontal, r- rotating, v – vertical). The AR is slightly different from tooth to tooth, best for incisors and premolars at the buccally smooth surfaces and rather low for the the canine 43 after horizontal brushing movement and molar 47 after horizontal brushing movement.

Agreement Rate Dr. Best Interdent clin with Dr. Best Interdent APP in % Delta (0)= AR (100%) in view of Tab. 5a, 5b, 5c

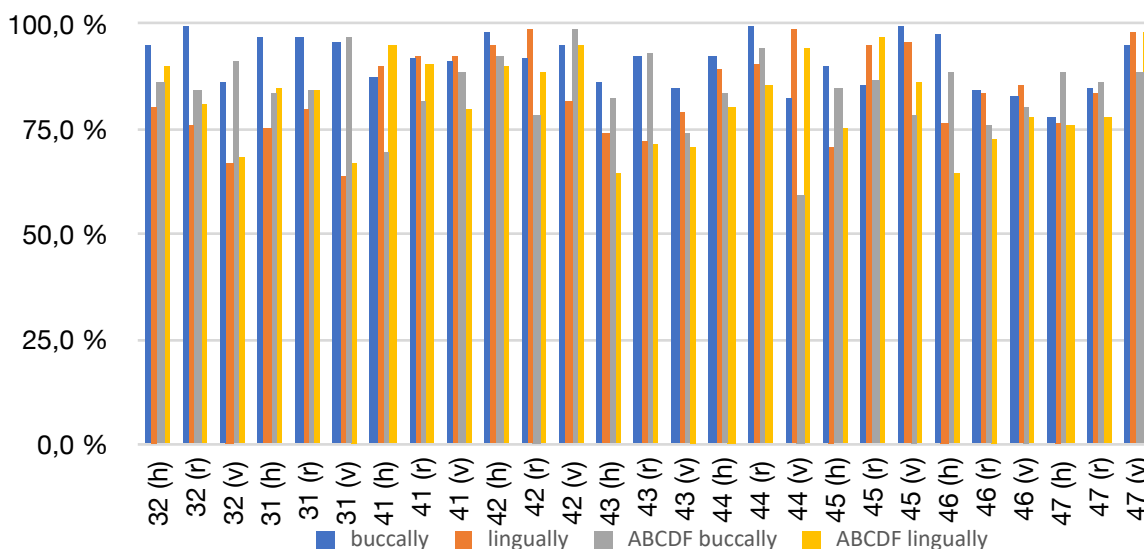


Fig. 14: Agreement Rate AR of toothbrushing with TB 2; AR per single incisors, canines, premolars and molars (h – horizontal, r- rotating, v – vertical). The AR is slightly different from tooth to tooth, best for the 4th quadrant incisors and premolars at the buccally smooth surfaces and rather low for the incisors in 3rd quadrant and 44 after vertical brushing movement at the ABCDF buccal risk field.

Agreement Rate clin with APP in % Delta (0)= AR (100%) in view of Fig. 9-11, Tab 4a, 4b and Tab 5a, 5b, 5c

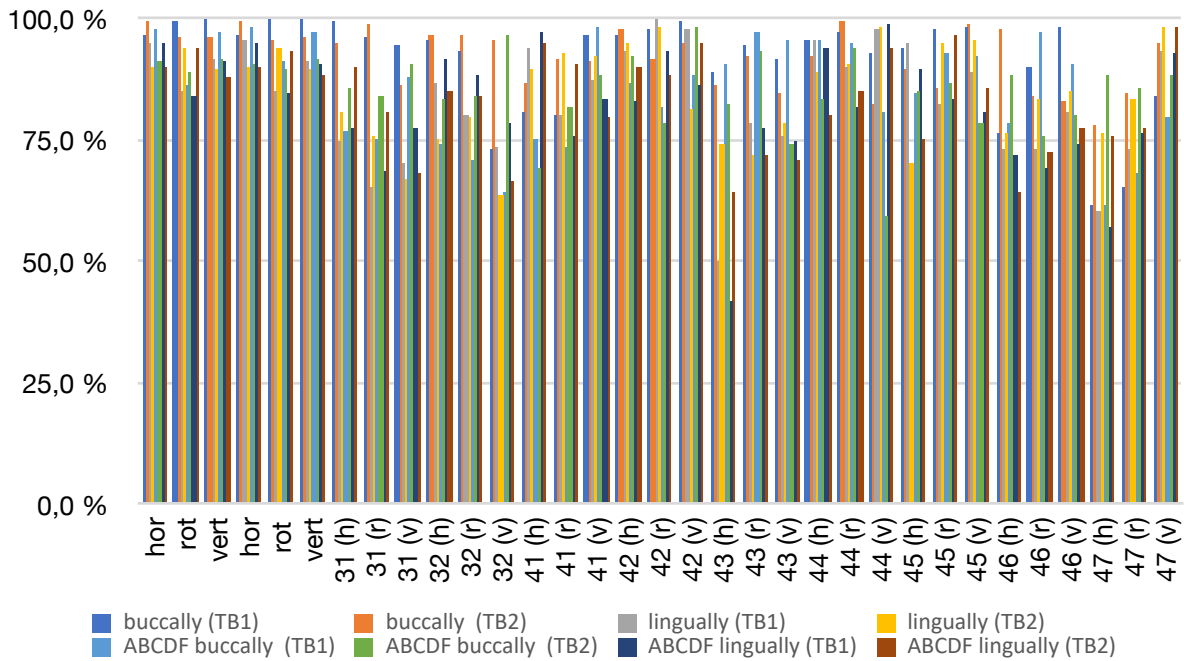


Figure 15: Agreement Rate AR of toothbrushing with TB 1 and TB 2; AR per single incisors, canines, premolars and molars (h - horizontal, r - rotating, v - vertically). The Agreement Rate is slightly different from tooth to tooth, best for incisors and premolars of the 4th quadrant and rather low for the canine 43.

Agreement Rate clin with APP in % Delta (0)= AR (100%) in view of Tab. 5a, 5b, 5c

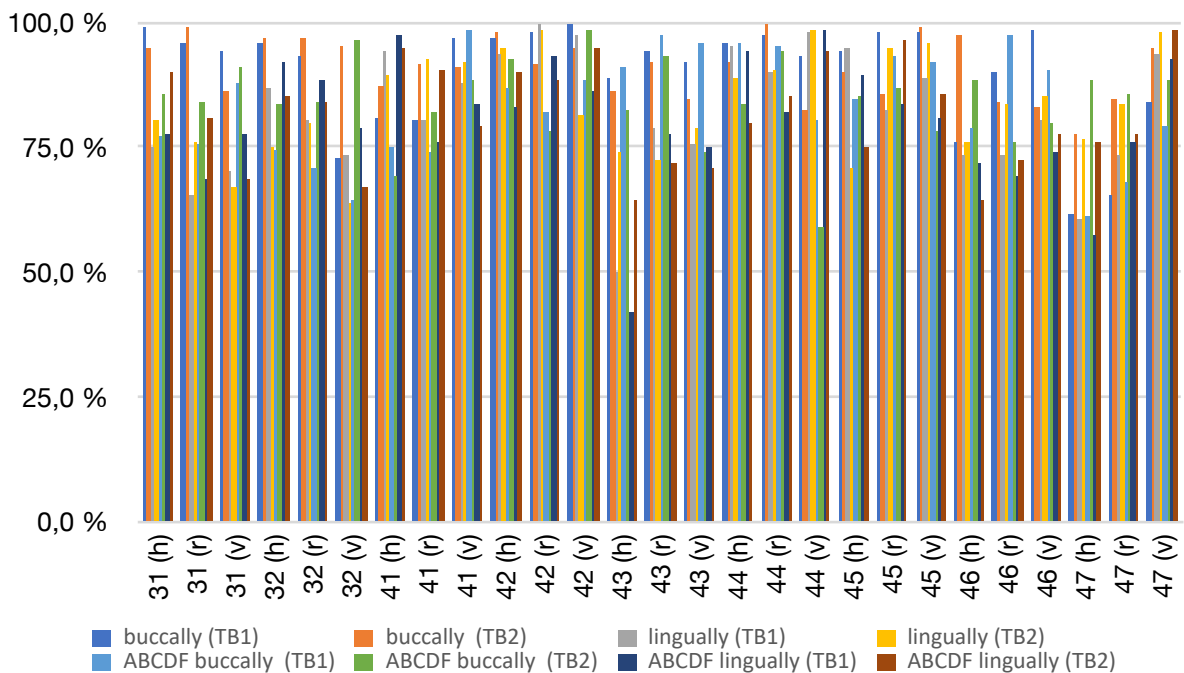


Figure 16: Agreement Rate AR of toothbrushing with TB 1 and TB 2; AR per single incisors, canines, premolars and molars (h - horizontal, r - rotating, v - vertically). The Agreement Rate is slightly different from tooth to tooth, best for incisors and premolars of the 4th quadrant and rather low for the canine 43.

4 Statistical analysis

4.1 Kolmogorov-Smirnov-Test

The K-S-test can be used to check whether there is a normal distribution of data or not.

The null hypothesis is defined as follows: H_0 = there is a normal distribution

A population of the observed data of $N = 18$ is available testing for each area. On the basis of the asymptomatic significance value corrected by Lilliefors (p (asymptomatic)) and the exact level of significance (p (exact)) it can be seen that **the assumption of the existence of a normal distribution is confirmed**. All calculated values are above the significance level, which was determined before the examination and indicated the probability with a high of $\alpha = 0.05$ (5%) with which an 1st kind of error occurs.

Tooth surface	N	Test statistic	p (asymptotic)	p (exact)
Buccally (APP)	18	0.190	.084 ^c	0.475
Lingually (APP)	18	0.099	.200 ^c	0.987
ABCDF Buccally (APP)	18	0.133	.200 ^c	0.865
ABCDF Lingually (APP)	18	0.088	.200 ^c	0.997
Buccally (clinical)	18	0.157	.200 ^c	0.709
Lingually (clinical)	18	0.117	.200 ^c	0.941
ABCDF Buccally (clinical)	18	0.107	.200 ^c	0.972
ABCDF Lingually (clinical)	18	0.109	.200 ^c	0.968

Tab. 1: Kolmogorov-Smirnov-Test of normal distribution, equality of means of plaque removal between robot program and clinical program, number of observations: N=18, significance value (p), corrected significance value by Lilliefors (C)

4.2 Independent t-Test

The independent t- test is used to assess whether the robot parameters match the clinical parameters.

Toothbrush / Brushing movement	Tooth surface	t	df	p	Mean difference	se
Flat Cut / Horizontal	Buccally	-0.710	4	0.517	-0.585	0.824
	Lingually	-1.089	4	0.337	-0.846	0.777
	ABCDF Buccally	0.438	4	0.684	0.188	0.430
	ABCDF Lingually	-1.241	4	0.282	-0.502	0.405
Flat Cut / Rotating	Buccally	-0.057	4	0.957	-0.041	0.710
	Lingually	-3.217	2.056	0.082	-2.686	0.835
	ABCDF Buccally	1.745	4	0.156	0.856	0.491
	ABCDF Lingually	-3.586	2.221	0.060	-1.532	0.427
Flat Cut / Vertical	Buccally	0.047	4	0.965	0.047	1.002
	Lingually	-1.574	2.395	0.236	-1.537	0.976
	ABCDF Buccally	0.365	4	0.734	0.263	0.721
	ABCDF Lingually	-1.542	2.447	0.240	-0.948	0.614
Interdental / Horizontal	Buccally	-0.129	2.021	0.909	-0.115	0.889
	Lingually	-1.781	2.005	0.217	-1.815	1.019
	ABCDF Buccally	1.579	4	0.189	0.944	0.598
	ABCDF Lingually	-1.768	2.126	0.212	-0.991	0.560
Interdental / Rotating	Buccally	1.096	4	0.335	0.710	0.648
	Lingually	-0.733	2.150	0.535	-1.114	1.521
	ABCDF Buccally	1.701	2.100	0.225	1.065	0.626
	ABCDF Lingually	-0.571	2.050	0.624	-0.657	1.151
Interdental / Vertical	Buccally	-1.553	4	0.195	-0.731	0.471
	Lingually	-1.599	4	0.185	-1.870	1.169
	ABCDF Buccally	-3.244	2.646	0.057	-0.843	0.260
	ABCDF Lingually	-2.021	2.205	0.169	-1.173	0.580

Tab. 2: Independent t-test for equality of means of plaque removal (assessed with PPI) between robot programme and clinical programme – separated by toothbrushes and brushing movements; test statistic of planimetric t-Test (t), degrees of freedom (df), significance value (p), difference of means of observations between robot program and clinical program, standard error of the mean difference (se). The independent t-test of means demonstrates the equality of all plaque removal parameters.

The independent t-test is used to assess whether the robot parameters match the clinical parameters. The null hypothesis is considered the agreement hypothesis: $H_0 : \mu_1 = \mu_2$. For example, the smaller the mean difference, the higher the correspondence of two mean values (μ). See the following ranks of mean difference in table 2a.

rank	Mean difference
1	Flat Cut / Rotating / buccally (-0.041)
2	Flat Cut / Vertical / buccally (0.047)
3	Interdental / Horizontal / buccally (-0.115)
4	Flat Cut / Horizontal / ABCDF buccally (0.188)
5	Flat Cut / Vertical / ABCDF buccally (0.263)
6	Flat Cut / Horizontal / ABCDF lingually (-0.502)
7	Flat Cut / Horizontal / buccally (-0.585)
8	Interdental / Rotating / ABCDF lingually (-0.657)
9	Interdental / Rotating / buccally (0.710)
10	Interdental / Vertical / buccally (-0.731)

Tab. 2a: Ranking table of the mean difference in relation to the independent t-test for equality of means of plaque removal (assessed with PPI) between robot programme and clinical programme – separated by toothbrushes and brushing movements – the mean difference of independent t-test demonstrates the equality of means to clarify the small mean differences, which the number in brackets represents. The order of the ranking table is from place 1 (lowest value) in ascending order to place 10 (higher value).

23 of 24 tests show a mean difference in the comparison between robot data and clinical data of less than 2. Especially the test Flat Cut / Rotating / buccally with mean difference of -0.041 and Flat Cut / Vertical / buccally with mean difference of 0.047 represents extremely low values. The only exception that achieved a higher difference with -2.686 is Flat Cut / Rotating / ABCDF lingually.

But not only these values say something about the agreement. The standard deviation rep. standard error (se) is also decisive in order to assess how significant the value of the mean difference is. The following ranks see table 2b.

rank	se
1	Interdental / Vertical / ABCDF buccally (0.260)
2	Flat Cut / Horizontal / ABCDF lingually (0.405)
3	Flat Cut / Rotating / ABCDF lingually (0.427)
4	Flat Cut / Horizontal / ABCDF buccally (0.430)
5	Interdental / Vertical / buccally (0.471)
6	Flat Cut / Rotating / ABCDF lingually (0.491)
7	Interdental / Horizontal / ABCDF lingually (0.560)
8	Interdental / Vertical / ABCDF lingually (0.580)
9	Interdental / Horizontal / ABCDF buccally (0.598)
10	Flat Cut / Vertical / ABCDF lingually (0.614)

Tab. 2b: Ranking table of the standard deviation (se) of the mean difference in relation to the independent t- test test for equality of means of plaque removal (assessed with PPI) between robot programme and clinical programme – separated by toothbrushes and brushing movements – the standard deviation shows how far on the average the values are from the mean difference. The corresponding amount is in brackets. In order to the ranking table is from place 1 (lowest se value) to place 10 (higher value).

In combination of mean difference and standard deviation, Flat Cut / Horizontal / ABCDF buccally with mean difference of 0.188 and standard deviation of 0.430 shows the highest agreement value. Further meaningful test, such as Flat Cut / Horizontal / ABCDF lingually shows a mean difference of -0.502 and standard deviation of 0.405 and Interdent / Vertical / buccally a mean difference of -0.731 and standard deviation of 0.471.

19 out of 24 data show a standard deviation of less than 1. Only 5 tests have a standard deviation between 1 and 1.6.

The independent t-test also uses the significance value (p) to show that the null hypothesis is confirmed in all tests. If $p > \alpha$, H_0 is considered as confirmed. If $p < \alpha$ the match hypothesis is rejected. The significance value is given as $\alpha = 0.05$. The following ranks see table 2c.

rank	p
1	Interdental / Vertical / ABCDF buccally (0.057)
2	Flat Cut / Rotating / ABCDF lingually (0.060)
3	Flat Cut / Rotating / lingually (0.082)
4	Flat Cut / Rotating / ABCDF buccally (0.156)
5	Interdental / Vertical / ABCDF lingually (0.169)
6	Interdental / Vertical / lingually (0.185)
7	Interdental / Horizontal / ABCDF buccally (0.189)
8	Interdental / Vertical/ buccal (0.195)
9	Interdental / Horizontal/ ABCDF lingually (0.212)
10	Interdental / Horizontal / lingually (0.217)
...	...
14	Flat Cut / Horizontal / ABCDF lingually (0.282)
15	Interdental / Rotating / buccally (0.335)
16	Flat Cut / Horizontal / lingually (0.337)
17	Flat Cut / Horizontal / buccally (0.517)
18	Interdental / Rotating / lingually (0.535)
19	Interdental / Rotating / ABCDF lingually (0.624)
20	Flat Cut / Horizontal / ABCDF buccally (0.684)
21	Flat Cut / Vertical / ABCDF buccally (0.734)
22	Interdental / Horizontal / buccally (0.909)
23	Flat Cut / Rotating / buccally (0.957)
24	Flat Cut / Vertical / buccally (0.965)

Tab. 2c: Ranking table of the significance value (p) difference in relation to the independent t- test test for equality of means of plaque removal (assessed with PPI) between robot programme and clinical programme – separated by toothbrushes and brushing movements – the significance value (p) indicated whether the test result is significant and whether the null hypothesis can be confirmed or not. The corresponding amount for the corresponding toothbrush, cleaning movement and the cleaning surfaces is shown in brackets. In order of the ranking table is from place 1 (lowest significance value) in ascending order to place 10 (higher value), and further in place 14 (even higher value) ascending to place 24 (highest significance value).

All values exceed previously defined level of significance value $\alpha = 0.05$ (5%). The higher the data deviation from α , the more significant the agreement hypothesis is. This is particularly true for Flat Cut / Vertical / buccally with $p = 0.965$ and Flat Cut / Rotating / buccally with 0.957.

The test variable t combines the data of sample into a value that is suitable for making a decision about the validity resp. significance of the null hypothesis.

After its calculation, t is compared with the critical t -value, which can be found in a table. If the t -value exceeds the critical value ($t_{df;1-(\alpha/2)}$), t is considered significance and the null hypothesis is confirmed.

Since a total of 6 values are usually compared with each other ($n_1 = 3, n_2 = 3$) the degrees of freedom are the number 4 ($df = n_1 + n_2 - 2$) for most data. Due to increases discrepancies in the following areas the degrees of freedom (df) were reduced by the decreased observations to 2 to 3. Nevertheless, with help of the test variables (t) and the significance value (p), it can be seen that even with these data recorded, significance and thus an assessment of the validity and the null hypothesis is guaranteed (table 2d).

df	p
Flat Cut / Rotating / lingually (2.056)	0.082
Flat Cut / Rotating / ABCDF lingually (2.221)	0.060
Flat Cut / Vertical / lingually (2.395)	0.236
Flat Cut / Vertical / ABCDF lingually (2.447)	0.240
Interdental / Horizontal / buccally (2.021)	0.909
Interdental / Horizontal / lingually (0.2005)	0.217
Interdental / Horizontal / ABCDF lingually (2.126)	0.212
Interdental / Rotating / ling (2.150)	0.535
Interdental / Rotating / ABCDF buccally (2.1)	0.225
Interdental / Rotating / ABCDF lingually (2.05)	0.624
Interdental / Vertical / ABCDF buccally (2.646)	0.057
Interdental / Vertical / ABCDF lingually (2.205)	0.169

Tab. 2d: The values are taken from independent t-test for equality of means of plaque removal (assessed with PPI) between robot programme and clinical programme – separated by toothbrushes and brushing movements. Test statistic of degrees of freedom (df) compared to the significance level (p). Representation of all df values < 4 with the corresponding significance level.

4.3 Wilcoxon-Mann-Whitney-U-Test

This W-M-W-U-Test is used to assess the central trend differences of an ordinary scaled feature, which in this case includes plaque removal. In this way all data of the respective groups (robot program and clinical program) are assigned rank numbers.

The test variable U indicates the frequency of rank places that are larger than the ranking places in the other observation group.

Toothbrush / Brushing movement	Tooth surface	U	Z	exact p
Flat Cut / Horizontal	Buccally	2.000	-1.091	0.400
	Lingually	2.000	-1.091	0.400
	ABCDF Buccally	3.000	-0.655	0.700
	ABCDF Lingually	2.500	-0.886	0.500
Flat Cut / Rotating	Buccally	4.000	-0.218	1.000
	Lingually	0.000	-1.964	0.100
	ABCDF Buccally	1.000	-1.528	0.200
	ABCDF Lingually	0.000	-1.964	0.100
Flat Cut / Vertical	Buccally	4.000	-0.221	1.000
	Lingually	0.500	-1.771	0.200
	ABCDF Buccally	4.000	-0.218	1.000
	ABCDF Lingually	0.500	-1.664	0.200
Interdental / Horizontal	Buccally	3.000	-0.655	0.700
	Lingually	1.000	-1.593	0.200
	ABCDF Buccally	2.000	-1.091	0.400
	ABCDF Lingually	0.500	-1.764	0.200
Interdental / Rotating	Buccally	3.000	-0.655	0.700
	Lingually	3.000	-0.655	0.700
	ABCDF Buccally	2.000	-1.091	0.400
	ABCDF Lingually	3.000	-0.655	0.700
Interdental / Vertical	Buccally	1.000	-1.528	0.200
	Lingually	2.000	-1.091	0.400
	ABCDF Buccally	0.000	-1.964	0.100
	ABCDF Lingually	0.500	-1.764	0.200

Tab. 3: Wilcoxon-Mann-Whitney-U-Test of equality of medians/rank sums of plaque removal (assessed with PPI) between robot program and clinical program – separated by toothbrushes and brushing movements, test statistic of non-parametric Mann-Whitney-Test (U), normalized test statistic (Z), significance value (p). The W-M-W-U-Test of means demonstrates the equality of all plaque removal parameters.

Z is used to check the statistical significance of these values and is compared with a critical z-value, which is dependent on the level of significance and is shown on a special table.

Using the specification of “exact p”, a significance check can be carried out using a table. If “exact p” > α the agreement hypothesis of plaque removal between the robot and clinical programme is confirmed. This is the case of all data.

Only 3 of 24 would decline the null hypothesis, if a higher significance level $\alpha = 0.10$ (10%) were used:

Toothbrush / Movement / Surface	exact p
Flat Cut / Rotating / lingually	0.100
Flat Cut / Rotating / ABCDF lingually	0.100
Interdental / Vertical / ABCDF buccally	0.100

Tab. 3a: Presentation of the exact level of significance (exact p) of the W-M-W-U-test - separated by toothbrushes and brushing movements - which would argue against the null hypothesis at a defined significance value of $\alpha = 0.1$ (10%).

5 Conclusions

The clinical validation of PG-Plaque simulation is based on the clinical test with two different manual toothbrushes with 22 video supported subjects compared to the robot test after meticulous transfer of clinical separated horizontal, rotating and vertical brushing movements.

The statistics are based on the clinical planimetric plaque index (clinPPI, Lang et al. 2011). Due to the high Agreement Rate AR of minimal 84% - 100% and to the independent t-test of equality of means as well as to the Wilcoxon-Mann-Whitney-U-test of equality of medians, the robot test programme is clinically validated in all three brushing movements.

6 References

Pepin H, Lang T, Weich K, Gaengler P (2020): Validation report, Clinical validation of Organic Plaque Simulation (PG Plaque) in Robot Toothbrushing Tests. ORMED, Germany: Journal of Dental Research, 2020, Vol. 99, Spec. Issue B Abstract No. 3075

Lang T, Stauffer S, Jennes B and Gaengler P (2014): Clinical validation of robot simulation of toothbrushing--comparative plaque removal efficacy. BMC Oral Health, 2014, 4;14:82

Altman DG (2016). Practical Statistics for Medical Research. London, UK: Chapman&Hall/CRC.

Barton B, Peat J (2014). Medical Statistics: A Guide to SPSS, Data Analysis and Critical Appraisal. Oxford, UK: Wiley Blackwell.

Bortz J, Schuster C (2016): Statistik für Human- und Sozialwissenschaftler. Springer-Lehrbuch. 7. Auflage, Sonderausgabe. Berlin: Springer

Bühl A (2016): SPSS 23. Einführung in die moderne Datenanalyse. Hallbergmoos: Pearson

7 Appendix

7.1 Single values of clinical testing

Tooth surface	Statistic	Brushing movement					
		Horizontal		Rotating		Vertical	
		Flat Cut	Inter-dental	Flat Cut	Inter-dental	Flat Cut	Inter-dental
Buccally (clinical)	M	7.40	6.45	6.67	6.73	7.58	8.40
	SD	1.01	1.54	0.94	1.05	1.71	0.76
	Med	7.78	6.62	6.46	6.67	7.49	8.25
	IQR LL	6.26	4.83	5.85	5.72	5.93	7.72
	IQR UL	8.17	7.89	7.70	7.81	9.33	9.22
Lingually (clinical)	M	7.88	8.89	9.32	8.93	9.43	9.91
	SD	0.82	1.76	1.44	2.59	1.61	1.91
	Med	7.56	8.22	9.59	9.89	8.67	10.22
	IQR LL	7.28	7.56	7.76	6.00	8.33	7.86
	IQR UL	8.82	10.89	10.59	10.90	11.28	11.64
ABCDF Buccally (clinical)	M	5.00	4.17	4.44	4.79	5.11	5.92
	SD	0.48	0.99	0.63	1.07	1.22	0.42
	Med	5.11	4.03	4.62	4.50	5.38	5.92
	IQR LL	4.47	3.25	3.74	3.89	3.78	5.50
	IQR UL	5.41	5.22	4.96	5.97	6.17	6.33
ABCDF Lingually (clinical)	M	5.54	6.14	6.57	6.21	6.54	6.95
	SD	0.46	0.96	0.72	1.98	1.01	0.98
	Med	5.44	5.78	6.89	7.09	6.00	6.61
	IQR LL	5.13	5.42	5.74	3.94	5.92	6.19
	IQR UL	6.04	7.22	7.07	7.60	7.70	8.06

Tab. A1: Mean (M), Standard deviations (SD), Median (Med) and Interquartile range (IQR, LL = lower limit, UL = upper limit) of post-brush parameters (PPI) clin for the toothbrushes Flat Cut, Interdent; separated for three brushing movements hor, rot, vert (Dr. Karl Weich, 2019, Robot Test of Cleaning Efficacy by Plaque Planimetry Validation of organic plaque 2019-07, Chapter 2.1)

7.2 Single values of robot testing

Tooth surface	Statistic	Brushing movement					
		Horizontal		Rotating		Vertical	
		Flat Cut	Inter-dental	Flat Cut	Inter-dental	Flat Cut	Inter-dental
Buccally (APP)	M	6.81	6.33	6.63	7.44	7.63	7.67
	SD	1.01	0.11	0.79	0.40	0.32	0.29
	Med	7.22	6.33	6.78	7.56	7.44	7.56
	IQR LL	5.67	6.22	5.78	7.00	7.44	7.44
	IQR UL	7.56	6.44	7.33	7.78	8.00	8.00
Lingually (APP)	M	7.04	7.07	6.63	7.81	7.89	8.04
	SD	1.07	0.06	0.17	0.50	0.51	0.68
	Med	7.22	7.11	6.67	7.78	8.00	7.89
	IQR LL	5.89	7.00	6.44	7.33	7.33	7.44
	IQR UL	8.00	7.11	6.78	8.33	8.33	8.78
ABCDF Buccally (APP)	M	5.19	5.11	5.30	5.85	5.37	5.07
	SD	0.57	0.29	0.57	0.17	0.28	0.17
	Med	5.33	5.00	5.44	5.89	5.33	5.11
	IQR LL	4.56	4.89	4.67	5.67	5.11	4.89
	IQR UL	5.67	5.44	5.78	6.00	5.67	5.22
ABCDF Lingually (APP)	M	5.04	5.15	5.04	5.56	5.59	5.78
	SD	0.53	0.17	0.17	0.22	0.34	0.22
	Med	5.22	5.11	5.00	5.56	5.67	5.78
	IQR LL	4.44	5.00	4.89	5.33	5.22	5.56
	IQR UL	5.44	5.33	5.22	5.78	5.89	6.00

Tab. A2: Mean (M), Standard deviations (SD), Median (Med) and Interquartile range (IQR, LL = lower limit, UL = upper limit) of post-brush parameters (PPI) APP for the toothbrushes Flat Cut, Interdent; separated for three brushing movements hor, rot, vert ((Dr. Karl Weich, 2019, Robot Test of Cleaning Efficacy by Plaque Planimetry Validation of organic plaque 2019-07; Chapter 2.1)

7.3 Single values tooth by tooth comparing clinical and robot testing

Tooth	Tooth surface	Brushing movement					
		Horizontal		Rotating		Vertical	
		Flat Cut	Inter-dental	Flat Cut	Inter-dental	Flat Cut	Inter-dental
31	Buccal (clin.)	6.13	6.27	5.62	7.17	7.67	9.17
	Lingual (clin.)	8.55	8.17	9.89	8.67	9.71	9.92
	ABCDF Buccally (clin.)	3.70	3.58	3.22	4.42	4.78	5.58
	ABCDF Lingually (clin.)	5.92	5.33	6.82	5.58	6.59	6.50
	Buccal (APP)	6.00	5.33	6.33	7.00	6.67	6.67
	Lingual (APP)	4.00	4.67	3.67	4.33	4.33	4.00
	ABCDF Buccally (APP)	6.00	5.00	5.67	6.00	6.00	4.67
	ABCDF Lingually (APP)	3.67	4.33	3.67	3.67	4.33	3.33
32	Buccal (clin.)	5.53	5.88	4.78	6.92	4.47	7.50
	Lingual (clin.)	7.02	8.48	7.91	9.33	8.79	10.50
	ABCDF Buccally (clin.)	3.42	3.67	2.73	4.08	3.12	4.33
	ABCDF Lingually (clin.)	4.80	5.50	5.18	6.25	5.48	7.33
	Buccal (APP)	6.33	5.33	6.00	6.33	9.33	6.67
	Lingual (APP)	4.67	4.00	4.33	5.67	4.00	4.00
	ABCDF Buccally (APP)	6.00	5.33	5.67	5.67	6.67	4.67
	ABCDF Lingually (APP)	4.00	4.00	4.00	4.67	3.33	4.00
41	Buccal (clin.)	5.90	6.33	4.78	6.83	6.92	8.75
	Lingual (clin.)	6.92	8.13	9.22	8.00	8.23	9.08
	ABCDF Buccally (clin.)	3.52	3.25	2.71	4.50	4.17	5.50
	ABCDF Lingually (clin.)	4.93	5.50	6.40	5.28	5.65	6.72
	Buccal (APP)	9.33	8.67	8.33	8.33	6.33	10.33
	Lingual (APP)	8.00	10.00	5.67	9.33	6.00	7.67
	ABCDF Buccally (APP)	6.00	6.33	5.33	6.33	4.00	6.67
	ABCDF Lingually (APP)	4.67	6.00	4.00	5.67	4.00	4.67

Tab. A3: Analysis of single teeth; post-brush plaque (PPI clin, PPI APP); teeth 31 - 41

Tooth	Tooth surface	Brushing movement					
		Horizontal		Rotating		Vertical	
		Flat Cut	Inter-dental	Flat Cut	Inter-dental	Flat Cut	Inter-dental
42	Buccal (clin.)	5.42	5.40	5.29	5.50	6.06	7.25
	Lingual (clin.)	6.15	7.43	7.67	8.42	7.91	9.33
	ABCDF Buccally (clin.)	3.33	3.25	3.84	3.83	3.83	4.83
	ABCDF Lingually (clin.)	4.32	5.00	5.67	5.50	5.29	5.83
	Buccal (APP)	6.00	5.00	5.67	7.00	6.00	6.33
	Lingual (APP)	7.33	8.33	7.67	8.67	8.33	6.00
	ABCDF Buccally (APP)	4.67	4.00	5.67	6.00	5.00	4.67
	ABCDF Lingually (APP)	6.00	6.00	6.33	6.67	6.67	5.33
43	Buccal (clin.)	5.69	4.15	4.64	5.25	6.22	7.08
	Lingual (clin.)	5.31	8.67	8.82	8.00	9.29	9.50
	ABCDF Buccally (clin.)	3.42	2.92	3.42	4.00	3.90	4.92
	ABCDF Lingually (clin.)	3.49	5.75	6.42	6.17	6.47	6.08
	Buccal (APP)	7.33	6.67	5.67	6.67	7.67	4.33
	Lingual (APP)	14.33	13.33	12.67	13.00	13.67	13.33
	ABCDF Buccally (APP)	4.33	4.67	3.67	4.67	4.33	2.33
	ABCDF Lingually (APP)	9.33	9.33	8.67	9.00	9.00	9.00
44	Buccal (clin.)	5.58	6.08	6.20	5.92	6.59	7.50
	Lingual (clin.)	7.17	9.63	9.44	10.03	9.96	10.08
	ABCDF Buccally (clin.)	4.75	4.33	4.18	4.58	4.95	6.42
	ABCDF Lingually (clin.)	5.58	7.00	6.82	6.81	7.13	7.58
	Buccal (APP)	6.33	4.67	6.67	6.00	5.33	4.33
	Lingual (APP)	8.00	7.67	7.67	8.33	10.33	10.33
	ABCDF Buccally (APP)	4.33	2.67	4.67	4.00	3.00	2.33
	ABCDF Lingually (APP)	5.00	5.00	5.00	5.33	7.00	7.00

Tab. A4: Analysis of single teeth; post-brush plaque (PPI clin, PPI APP); teeth 42 - 44

Tooth	Tooth surface	Brushing movement					
		Horizontal		Rotating		Vertical	
		Flat Cut	Inter-dental	Flat Cut	Inter-dental	Flat Cut	Inter-dental
45	Buccal (clin.)	8.08	5.83	6.96	6.08	7.34	8.17
	Lingual (clin.)	7.57	10.32	9.53	8.92	9.96	9.42
	ABCDF Buccally (clin.)	6.85	4.83	5.29	5.33	5.77	6.83
	ABCDF Lingually (clin.)	6.05	7.50	7.00	6.67	7.27	7.42
	Buccal (APP)	7.00	7.67	7.33	8.67	7.00	8.00
	Lingual (APP)	6.67	5.00	6.33	8.00	8.00	8.67
	ABCDF Buccally (APP)	5.33	6.33	6.00	6.67	5.00	4.67
	ABCDF Lingually (APP)	5.00	4.00	4.33	6.33	5.33	6.00
46	Buccal (clin.)	10.95	7.42	9.82	6.83	10.08	8.25
	Lingual (clin.)	10.13	9.95	9.82	9.00	9.83	10.33
	ABCDF Buccally (clin.)	7.12	4.83	6.42	4.92	6.73	6.33
	ABCDF Lingually (clin.)	6.83	7.25	7.09	6.75	7.59	7.58
	Buccal (APP)	6.67	7.00	8.00	9.67	10.33	11.33
	Lingual (APP)	5.33	5.67	5.00	6.00	6.33	7.67
	ABCDF Buccally (APP)	5.00	6.00	6.67	7.33	7.67	8.33
	ABCDF Lingually (APP)	4.00	3.67	4.00	4.00	4.00	5.33
47	Buccal (clin.)	13.31	10.67	11.94	10.11	12.88	11.92
	Lingual (clin.)	12.13	9.22	11.53	10.00	11.17	11.00
	ABCDF Buccally (clin.)	8.87	6.83	8.14	7.42	8.72	8.50
	ABCDF Lingually (clin.)	7.93	6.42	7.72	6.92	7.39	7.50
	Buccal (APP)	6.33	6.67	5.67	7.33	10.00	11.00
	Lingual (APP)	5.00	5.00	6.67	7.00	10.00	10.67
	ABCDF Buccally (APP)	5.00	5.67	4.33	6.00	6.67	7.33
	ABCDF Lingually (APP)	3.67	4.00	5.33	4.67	6.67	7.33

Tab. A5: Analysis of single teeth; post-brush plaque (PPI clin, PPI APP); teeth 45 - 47