Title Page

Title: The salivary oxidative stress profile shifts in paediatric dentofacial infection: a multivariate analysis of nine markers

Author list:

Mateusz F. Kolek*[†] (https://orcid.org/0000-0001-6470-4830)
Center for Cancer Research, Medical University of Vienna, Borschkegasse 8a, 1090 Vienna, Austria

Robert Dłucik
Dlucik Dental Clinic, 40-737 Katowice, Children's Maxillofacial Surgery Clinic, Chair of Pediatric Surgery, Medical University of Silesia, Medyków 16, 40-752 Katowice, Poland.

Jolanta Zalejska-Fiolka (https://orcid.org/0000-0002-3879-6548)
Department of Biochemistry, Faculty of Medical Sciences in Zabrze, Medical University of Silesia (SUM), Jordana 19, 41-800 Zabrze, Poland

Natalia Sitek Ignac (https://orcid.org/0000-0003-0550-4999)
Children's Maxillofacial Surgery Clinic, Chair of Pediatric Surgery, Medical University of Silesia, Medyków 16, 40-752 Katowice, Poland.

Bogusława Ewa Orzechowska-Wylęgała[†] (https://orcid.org/0000-0001-5950-2425)
Children's Maxillofacial Surgery Clinic, Chair of Pediatric Surgery, Medical University of Silesia, Medyków 16, 40-752 Katowice, Poland.

* - corresponding author (mateusz.kolek@meduniwien.ac.at)
[†] - these authors contributed equally

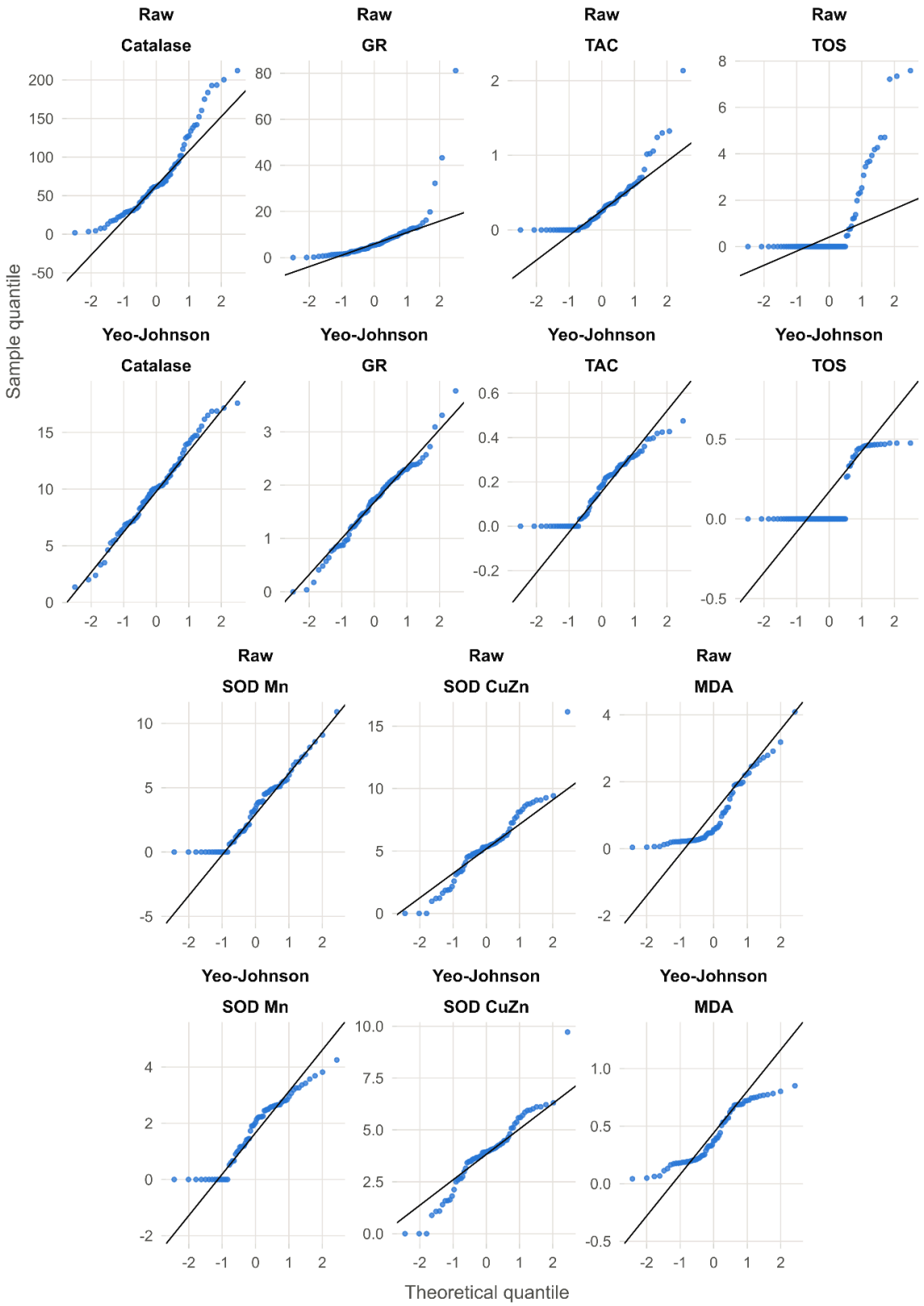


Fig. S1 Yeo-Johnson transformation The transform normalises the markers with continuous distributions and cannot normalise the zero-inflated ones, where a point mass at zero survives any monotone transformation. Normal quantile–quantile plots for the seven independent markers

before (top row) and after (bottom row) the Yeo-Johnson power transform, with one maximum-likelihood λ estimated per marker. Points are the observed values against the quantiles a normal distribution of the same mean and variance would produce; the diagonal is the reference line, so points lying along it indicate normality and systematic departures from it indicate skew, heavy tails or ties. The transform succeeds where the underlying distribution is continuous. Catalase straightens from skewness 1.00 to -0.04 and glutathione reductase from 4.88 to −0.006. These are the only two markers whose transformed values are not distinguishable from normal. Five of the seven markers remain formally non-normal after transformation. This does not invalidate the PCA, which assumes no distribution: the transform is applied to prevent extreme skewness and isolated large values from dominating the correlation matrix and the resulting components, and by that measure it succeeds for every marker - every skewness is reduced.

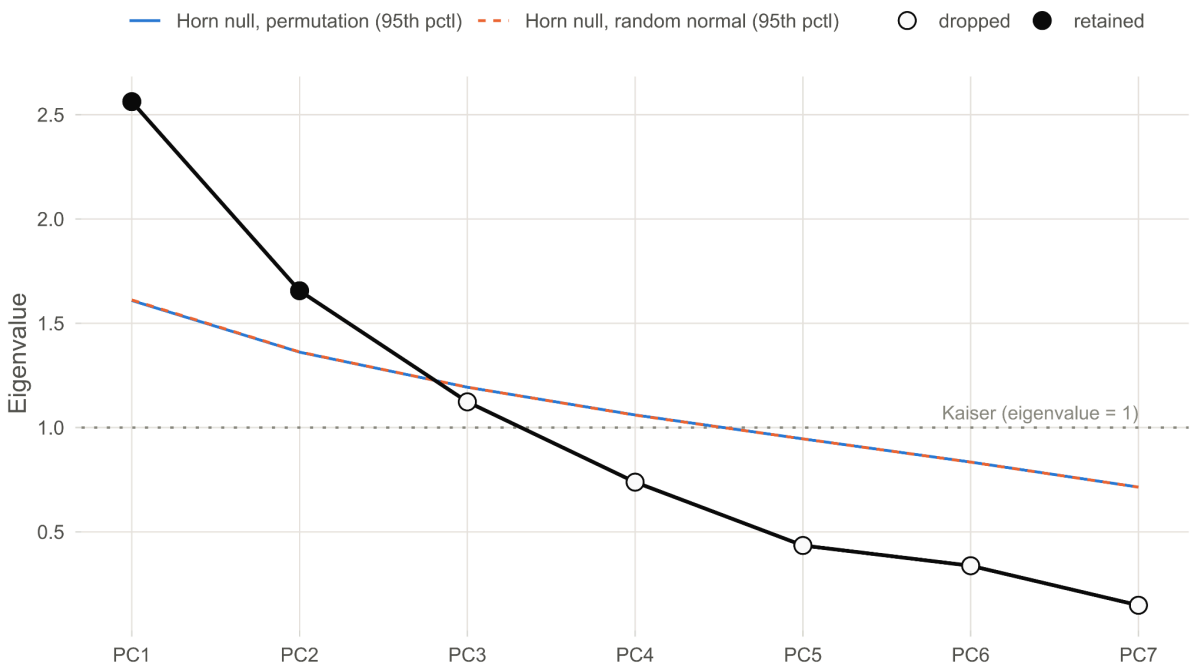


Fig. S2 Two components stand above the parallel-analysis null. The third exceeds Kaiser's criterion but not the Horn null and was dropped from the subsequent analyses. Scree plot of the seven eigenvalues from the correlation PCA of the seven independent markers (all 80 subjects, Yeo-Johnson transformed, missing values filled by regularised iterative PCA). Because the analysis is run on the correlation matrix, the eigenvalues sum to 7 and each is read in units of one marker's variance: a component with an eigenvalue of 1 carries as much variance as a single standardised marker. The black line traces the observed eigenvalues. Filled points are the two retained components, open points the five dropped. Three retention criteria are drawn against it. The dotted horizontal line is Kaiser's criterion (eigenvalue > 1), which retains three. The blue solid and orange dashed lines are the 95th percentiles of Horn's parallel analysis under two null models - 25 000 permutations of each marker's own values, and 25 000 draws of independent N(0,1) data of the same dimensions - each giving the largest eigenvalue that sampling variation alone would be expected to produce at that position in the sequence.

Horn's parallel analysis via permutation-the pre-specified retention method-retained two principal components. The first two components comfortably exceeded their upper-percentile null thresholds: PC1 (eigenvalue 2.563 vs. null 1.609) and PC2 (1.656 vs. 1.362). In contrast, PC3 (1.123) fell below its null expectation (1.193), with the observed scree curve intersecting the null boundary between PC2 and PC3. Kaiser's criterion (eigenvalue > 1) would retain PC3 only by failing to account for sampling variation; with seven variables, Kaiser's rule tends to over-retain because the expected chance eigenvalue at position three is 1.19 rather than 1.00. Retention of two components was robust across implementations: permutation and standard null models yielded visually indistinguishable thresholds, and an independent analysis (psych::fa.parallel) independently selected two

components. The acceleration factor also supported a two-component solution, showing a clear elbow at PC2 (0.374 vs. 0.148 at PC3). Together, PC1 and PC2 explain 36.6% and 23.7% of the total variance, respectively (60.3% cumulative). Adding PC3 would increase cumulative variance to 76.3%, but at the cost of retaining noise.

Table S1. Descriptive statistics and normality diagnostics for salivary oxidative-stress markers. Values are presented by group: Control (n = 32) and Disease (n = 48). Analyses are pairwise-complete; effective sample sizes (n) equal total group size minus missing observations (n_missing). Statistics include the interquartile range (IQR = Q3 - Q1) and type-2 skewness and excess kurtosis (R package *e1071*). Normality was evaluated using Shapiro-Wilk tests. Because normality was declared in only 5 of 18 cells (total SOD and CuZn-SOD in both groups; Mn-SOD and catalase in Controls), non-parametric methods were used for all downstream group comparisons. Units are specified in variable labels. All values, except from p-values are rounded to 2 decimal places. This table presents uncorrected raw distributions. Two control subjects had sub-floor total antioxidant capacity (TAC) readings (-0.00107 and -0.00226 mmol/g protein), which propagated into extreme negative oxidative stress index values (OSI = TOS /TAC / 10; minimum = -70.18). These two TAC and OSI values were censored as missing in all subsequent inferential analyses, though unaffected markers for these subjects (catalase, glutathione reductase, SOD, and MDA) were retained.

Variable	group	n total	n missing	n	mean	sd	median	iqr	min	max	skewness	kurtosis	Shapiro-Wilk's W	Shapiro-Wilk's p
Catalase (IU/g protein)	Control	32	0	32	73.77	42.26	66.30	51.24	1.99	183.61	0.82	0.39	0.95	0.133
	Disease	48	0	48	71.27	56.68	60.21	64.72	3.44	211.99	1.06	0.25	0.88	<0.001
Glutathione reductase (IU/g protein)	Control	32	0	32	9.04	13.85	7.03	6.76	0.04	81.12	4.81	25.45	0.45	<0.001
	Disease	48	0	48	6.78	7.86	5.20	6.87	0.00	43.30	3.00	11.06	0.68	<0.001
TAC (mmol/g protein)	Control	32	0	32	0.31	0.36	0.17	0.51	0.00	1.32	1.43	1.92	0.81	<0.001
	Disease	48	0	48	0.35	0.41	0.28	0.44	0.00	2.13	2.23	6.90	0.77	<0.001
TOS (µmol/g protein)	Control	32	1	31	0.91	1.86	0.00	0.76	0.00	7.34	2.23	4.36	0.56	<0.001
	Disease	48	0	48	0.96	1.85	0.00	1.00	0.00	7.58	2.22	4.70	0.59	<0.001
SOD total (NU/mg protein)	Control	32	6	26	9.57	4.05	10.14	4.06	0.00	17.24	-0.45	0.59	0.97	0.665
	Disease	48	5	43	7.97	3.87	7.88	4.51	0.00	19.80	0.46	1.33	0.97	0.250
SOD Mn (NU/mg protein)	Control	32	7	25	2.98	2.53	2.14	3.48	0.00	9.09	0.56	-0.30	0.92	0.062
	Disease	48	6	42	3.58	2.85	3.62	4.49	0.00	10.89	0.40	-0.52	0.94	0.019

SOD CuZn (NU/mg protein)	Control	32	6	26	6.70	3.18	6.38	3.63	0.00	16.14	0.43	2.42	0.92	0.053
	Disease	48	5	43	4.47	2.04	4.89	2.35	0.00	8.90	-0.38	-0.07	0.97	0.222
MDA (µmol/g protein)	Control	32	9	23	0.43	0.43	0.32	0.25	0.07	2.23	3.69	15.66	0.58	<0.001
	Disease	48	6	42	1.38	1.06	1.23	1.87	0.04	4.09	0.43	-0.68	0.93	0.010
OSI (AU)	Control	32	9	23	-2.68	14.73	0.00	0.46	-70.18	2.44	-4.78	22.89	0.25	<0.001
	Disease	48	11	37	0.42	0.87	0.00	0.45	0.00	4.40	3.12	11.80	0.55	<0.001

Table S2. Principal component analysis (PCA) of salivary oxidative-stress markers: Loadings, communalities, contributions, and sampling adequacy. PCA was conducted on seven primary markers (n = 80 pooled subjects). Two composite variables, total SOD (Mn-SOD + CuZn-SOD) and OSI (TOS/ TAC/10), were excluded to avoid artificial collinearity. This subset aligns with the PERMANOVA marker set. Variables were Yeo-Johnson transformed (see Supplementary figure S1) to reduce skewness, and missing values (0-18.75% per variable) were imputed via regularized iterative PCA (missMDA) with cross-validated rank selection. The PCA was performed on the correlation matrix of the transformed, imputed data (centered and scaled to unit variance). Two components were retained based on pre-specified permutation-based Horn's parallel analysis (Supplementary figure S2). Loadings (loading_PC1, loading_PC2) represent variable-component correlations; cos^2 denotes the squared loading. Communality (cos^2_PC1} + cos^2_PC2) reflects the proportion of variance captured by the two-component model; uniqueness equals 1 - communality. Percentage contributions (contrib_PC1_pct, contrib_PC2_pct) sum to 100% per component, with 14.3% (100/7) representing equal expected contribution. Measures of sampling adequacy (msa) include individual KMO values. Overall matrix KMO was 0.604 (mediocre), and Bartlett's test of sphericity confirmed suitability for factor analysis (chi^2 = 195.13, df = 21, $p < 0.001$). Rows are sorted by descending absolute PC1 loadings. All values are rounded to four decimal places.

Variable	loading_PC1	loading_PC2	communality	uniqueness	contrib_PC1_pct	contrib_PC2_pct	cos^2_PC1	cos^2_PC2	msa
Catalase	-0.8755	0.2378	0.8230	0.1770	29.9092	3.4148	0.7665	0.0565	0.6059
GR	-0.8243	0.1778	0.7110	0.2890	26.5111	1.9083	0.6794	0.0316	0.6296
MDA	0.7246	-0.1201	0.5394	0.4606	20.4858	0.8710	0.5250	0.0144	0.6389
SOD CuZn	-0.5966	-0.4897	0.5958	0.4042	13.8895	14.4825	0.3559	0.2398	0.7016
TAC	0.4776	0.4077	0.3944	0.6056	8.9020	10.0383	0.2281	0.1662	0.5707
TOS	-0.0803	0.8530	0.7340	0.2660	0.2518	43.9329	0.0065	0.7275	0.5151
SOD Mn	-0.0360	-0.6479	0.4211	0.5789	0.0506	25.3522	0.0013	0.4198	0.4439