

## Are You a P-Value Worshipper?

The trend of the over-emphasis on having the results of a research work to be statistically significant ( $P < .05$ ) is still going strong today due to the fact most researchers are statistically-phobiaed. In this write-up, I want to encourage a research paper reader to firstly critique on the research process. Table 1 shows the stages of a research study that need to be addressed in detail before a credible and clinically relevant result could be obtained.

It is essential that stages 1 & 2 be properly set-up (available, hopefully, in the Materials & Methods of a paper) otherwise, even with the help of a statistician the results obtained will not be valid!

For the results, the important question to ask is "Is the work clinically relevant to me?" An important point for a P-value worshipper to take note: "P-value is influenced by sample size, the larger the sample size, the likelihood of  $P < .05$  is increased!". For example, a researcher wants to determine the correlation between airway volume & lower face height; Table 2a shows a relatively poor correlation of  $r = 0.271$ ,  $P = 0.100$  with  $n = 38$ . But when  $n$  was doubled, though the relationship remains poor, the P-value has become significant ( $P = 0.018$ ), see Table 2b – the impact of sample size! Figure 1 shows the graphical presentation of the poor relationship. A good clinical relationship

(say between lower face height and anterior face height,  $r = 0.827$ ) will be given by  $r > 0.7$  (Figure 2).

In a correlation analysis, both variables are taken to be dependent. If we want to use lower face height to predict airway volume, the squaring of the correlation ( $r = 0.217$ ) shows that lower face height only explains about 5% of the variation in airway volume; whereas lower face height will explain 68% (squaring 0.827) of the variance in anterior face height. We will need the adjusted r-square of a multiple linear regression model to be high (at least 0.8) if we want to use the model for the prediction of the outcome variable. But if one is interested to determine significant predictors on the outcome variable, then the value of the adjusted r-square is not crucial in the interpretation anymore; since the interest is on the individual-predictor's P-value.

Table 3 shows the 4 combinations a research study can have on their clinical and statistical significances.

You are right! The "Clinical significance" should be focused first then the p-value. Scenarios 1 and 3 will be published but scenario 2 will miss a potential intervention as the possibility of getting a publication will be low because of  $P > .05$ !

For the statistically-phobiaed, Table 4 gives a summary of the various statistical techniques (the detailed discussions are given in references

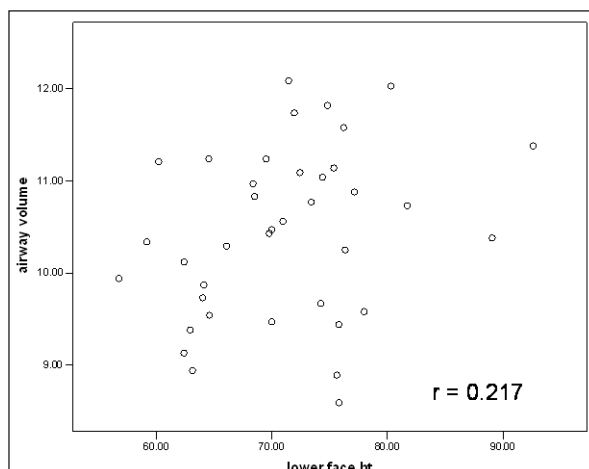


Figure 1. Scatter plot of a poor relationship.

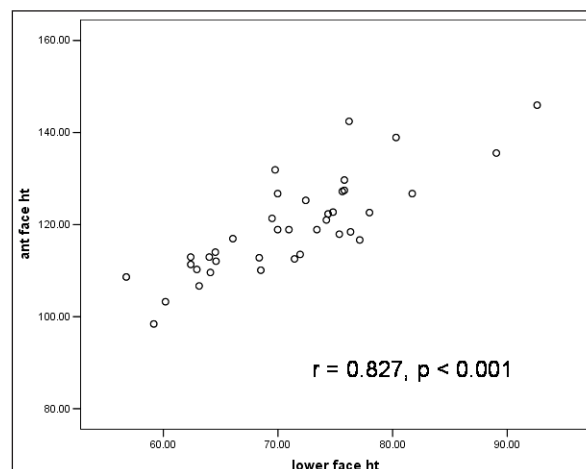


Figure 2. Scatter plot of a meaningful clinical relationship.

**Table 1.** Stages of a research process.<sup>1</sup>

Stages	Percentage of contribution to validity of clinical results obtained
Stage 1 Proper study design (Epidemiological / Randomized Controlled Trial) Sample size calculations <sup>2</sup> (Precision / Power calculations)	30 – 40%
Stage 2 Conduct of study / Data integrity (Garbage in Garbage out)	50 – 60%
Stage 3 Proper database setup / Statistics	10 – 20%

**Table 2a.** Correlation with n=38.

Correlations			
		Airway volume	Lower face height
Airway volume	Pearson correlation	1	.271
	Sig. (2-tailed)		.1
	N	38	38
Lower face height	Pearson correlation	.271	1
	Sig. (2-tailed)	.1	
	N	38	38

**Table 2b.** Correlation with n=76.

Correlations			
		Airway volume	Lower face height
Airway volume	Pearson correlation	1	.271*
	Sig. (2-tailed)		.018
	N	76	76
Lower face height	Pearson correlation	.271*	1
	Sig. (2-tailed)	.018	
	N	76	76

\*: Correlation is significant at the .05 level (2-tailed).

**Table 3.** Clinical vs statistical significance.

Clinical significance achieved?	Statistical significance achieved?	Implications of study results
Yes	Yes	Reproducible valuable study
Yes	No	Potential trend but could be due to chance (Pilot study)
No	Yes	Large sample size – still good information on 'best' clinical efficacy achieved.
No	No	Forget about it!

**Table 4.** Summary of statistical techniques.

Quantitative data		
Univariate		Multivariate
Normality assumed (Parametric)	Normality not assumed (Non-parametric)	
1 Sample T (e.g. Weight of today's infants compared to the reference-norm 5 years ago)	Wilcoxon Signed Rank	
Paired T (e.g. Pre-Post intervention weight loss)		
<b>Homogeneity (Equal variance) assumed</b>		N-way Anova (All independent variables are categorical)
Yes	No	Linear regression (All independent variables are quantitative)
2 Sample T (e.g. Differences in SBP between gender)	Mann Whitney U Wilcoxon Rank Sum	ANCOVA: Analysis of covariance (mixture of categorical and quantitative independent variables)
One-way Anova (e.g. Differences in SBP amongst BMI groups – underweight, normal weight, overweight, obese)	Kruskal Wallis	
Pearson's Correlation (e.g. Association between SBP and age)	Spearman's Correlation	
Qualitative data		
Test of associations (e.g. Smoking with cancer) Strength of associations	Chi-square (large n) Fisher's exact (small n) Relative risk (prospective outcomes) Odds ratios (non-prospective outcomes)	Logistic regression
Matched case-control Pre-Post qualitative outcomes	McNemar	Conditional logistic regression
Time to event data (Survival analysis)		
Kaplan Meier		Cox regression

3–9) that have a coverage of about 75–80% of all analyses performed in published articles; otherwise you may want to refer to the references 10–18 or alternatively seek a consult from a statistician.

In conclusion, statistics is akin to a oven in a cake-baking process; an essential apparatus but the quality of the cake predominantly depends on the baker (the researcher) and the quality of the ingredients (data quality), though the brand of the oven does enhance a better cake-quality. It is strongly encouraged to get a statistician involved in the planning stage of your study to assist in the Stages 1 & 2 of the research process before finally setting up the database and statistical analysis. Are you still a p-value worshipper? I wish - no more, hurray!

## REFERENCES

1. Chan YH. Randomized controlled trials (RCTs)–Essentials. *Singapore Med J* 2003;44:60–63.
2. Chan YH. Randomized controlled trials (RCTs)–Sample size: The magic number? *Singapore Med J* 2003;44:172–174.
3. Chan YH. Biostatistics 101: Data presentation. *Singapore Med J* 2003;44:280–285.
4. Chan YH. Biostatistics 102: Quantitative data: Parametric & non-parametric tests. *Singapore Med J* 2003;44:391–396.
5. Chan YH. Biostatistics 103: Qualitative data: Tests of independence. *Singapore Med J* 2003;44:498–503.
6. Chan YH. Biostatistics 104: Correlational analysis. *Singapore Med J* 2003;44:614–619.
7. Chan YH. Biostatistics 201: Linear regression analysis. *Singapore Med J* 2004;45:55–61.
8. Chan YH. Biostatistics 202: Logistic regression analysis. *Singapore Med J* 2004; 45:149–153.
9. Chan YH. Biostatistics 203: Survival analysis. *Singapore Med J* 2004;45:249–256.
10. Chan YH. Biostatistics 301: Repeated measurement analysis. *Singapore Med J* 2004;45:354–368.
11. Chan YH. Biostatistics 301a: Repeated measurement analysis (Mixed models). *Singapore Med J* 2004;45:456–460.
12. Chan YH. Biostatistics 302: Principal component and factor analysis. *Singapore Med J* 2004;45:558–566.
13. Chan YH. Biostatistics 303: Discriminant analysis. *Singapore Med J* 2005;46:54–61.
14. Chan YH. Biostatistics 304: Cluster analysis. *Singapore Med J* 2005;46:153–159.
15. Chan YH. Biostatistics 305: Multinomial logistic regression. *Singapore Med J* 2005;46:259–268.
16. Chan YH. Biostatistics 306: Log-linear models: poisson regression. *Singapore Med J* 2005;46: 377–385.
17. Chan YH. Biostatistics 307: Conjoint analysis and canonical correlation. *Singapore Med J* 2005;46:514–517. *Erratum in: Singapore Med J* 2005;46:740.
18. Chan YH. Biostatistics 308. Structural equation modeling. *Singapore Med J* 2005;46:675–679.