16

Comparison of Two Measures of Missing Cofactor in Cofactor-Dependent Enzymes: Proportion Versus Relative Increase

Tomas Hansson*

Department of Applied Environmental Science (ITM), Stockholm University, SE-10691 Stockholm, Sweden

Abstract: The degree of missing cofactor in a cofactor-dependent enzyme is widely used as a biomarker of cofactor deficiency. The degree of missing cofactor can be expressed either as the proportion of enzyme without cofactor, or as the relative increase in enzyme with cofactor after addition of excess cofactor to the sample. Especially for enzymes with thiamine pyrophosphate (TPP) as a cofactor, the relative increase (TPP-effect) has been used in a majority of studies, and its use seems to prevail without consideration of the proportion (latency) as a better alternative. In this letter, the statistical properties of the two measures are compared in the context of a thiamine-dependent enzyme. Proportion is a more balanced and sensitive measure than relative increase, and simulation shows that proportion is associated with equal or higher statistical power than relative increase. The power difference can be as high as 0.12.

Keywords: Cofactor, Enzyme, Latency, Statistical power, TPP-effect.

1. INTRODUCTION

For several decades, researchers measuring the degree of missing cofactor in cofactor-dependent enzymes have predominantly expressed their results as the relative increase in enzyme with cofactor after addition of excess cofactor to the sample. Especially for thiamine-dependent enzymes, this practice seems to prevail (see [1] for an example) without consideration of the proportion of enzyme without cofactor as a better alternative (see [2] for an exception to the rule). In this letter, the statistical properties of the two measures are compared in a context of the thiamine-dependent enzyme α -ketoglutarate dehydrogenase. The comparison is valid for cofactor-dependent enzymes in general, but only for parametric statistical methods. If non-parametric statistical methods are used, the two measures give identical results.

2. LATENCY AND TPP-EFFECT

Thiamine-dependent enzymes need TPP as a cofactor for their function. If the cofactor is missing, the thiaminedependent enzyme is inactive. The amount of enzyme lacking cofactor, commonly referred to as *apoenzyme*, is widely used as a biomarker of thiamine deficiency. The amount of apoenzyme is determined by measuring the activity of a thiamine-dependent enzyme both before and after addition of excess cofactor *in vitro*. The rationale of this procedure is that the supplied extra cofactor binds to the apoenzyme and thereby turns it into a fully functional *holoenzyme*. The obtained activities are referred to as the *endogenous activity* (obtained before addition of excess cofactor) and the *maximum activity* (obtained after addition of excess cofactor), respectively. The two activities are customarily assumed to be directly proportional to the amount of active enzyme. Hence, the original amount of apoenzyme is obtained as the difference between the maximum activity and the endogenous activity. This difference can be expressed in percent of the maximum activity (latency) or the endogenous activity (TPP-effect):

$$Latency = 100 \times \frac{Maximum \ activity - Endogenous \ activity}{Maximum \ activity}$$
$$TPP-effect = 100 \times \frac{Maximum \ activity - Endogenous \ activity}{Endogenous \ activity}$$

The latency is the proportion of apoenzyme, whereas the TPP-effect is the relative increase in the amount of active enzyme after excess cofactor addition. Theoretically, the maximum activity is always larger than or equal to the endogenous activity, which in turn is always larger than zero. Hence, the latency is always a number between 0 and 100, whereas the TPP-effect is positive and approaches infinity as the endogenous activity approaches zero. For a fixed maximum activity, the latency is a linear function of the endogenous activity, whereas the TPP-effect is a non-linear function of the endogenous activity. (For a fixed maximum activity, the TPP-effect is a linear function of the reciprocal of the endogenous activity.) Consequently, small and large latency values have equal influence on *e.g.* the arithmetic mean of a sample, whereas large TPP-effect values have higher influence than small TPP-effect values. In this respect, latency is a more balanced measure than TPP-effect. Moreover, the coefficient of variation of the latency is always smaller than or equal to the coefficient of variation of the TPP-effect, as shown mathematically in APPENDIX A. Hence, latency is a more sensitive measure than TPP-effect. The question may be raised, however, whether the differences in balance and sensitivity are of any practical importance, e.g. for the statistical power associated with the respective measure.

^{*}Address correspondence to this author at the Department of Applied Environmental Science (ITM), Stockholm University, SE-10691 Stockholm, Sweden; Tel: +46-8-6747368; Fax: +46-8-6747638; Email: tomas.hansson@itm.su.se

3. SIMULATION OF STATISTICAL POWER

The respective powers associated with latency and TPPeffect were investigated by testing, with Student's *t*-test, the null hypothesis (H₀) of equal means of two random samples drawn from populations with different means. The models were designed to mimic, as closely as possible, the endogenous and maximum activities of the thiamine-dependent enzyme α -ketoglutarate dehydrogenase in thiamine-deficient young herring gulls [3]. α -Ketoglutarate dehydrogenase catalyses a chemical reaction, in which one of the products is reduced nicotinamide adenine dinucleotide (NADH). The endogenous and maximum activities of this enzyme are expressed as nmol NADH formed per minute and mg protein. The simulation was made with the software Intercooled Stata 9.2 (StataCorp LP, College Station, TX, USA), and the program code is provided in **APPENDIX B**.

3.1. Random Variables for Two Populations *i* = 1, 2

 X_i = endogenous activity of a specimen randomly selected from population *i* [nmol/min/mg]

 Y_i = maximum activity of a specimen randomly selected from population *i* [nmol/min/mg]

 $W_i = 100 \cdot \frac{Y_i - X_i}{Y_i}$ = latency of a specimen randomly selected

from population *i* [%]

 $Z_i = 100 \cdot \frac{Y_i - X_i}{X_i}$ = TPP-effect of a specimen randomly se-

lected from population *i* [%]

3.2. Samples from Two Populations i = 1, 2

 $\mathbf{x}_i = (\mathbf{x}_{i1}, \mathbf{x}_{i2}, \dots, \mathbf{x}_{ij}) =$ sample containing *j* observations of X_i $\mathbf{y}_i = (\mathbf{y}_{i1}, \mathbf{y}_{i2}, \dots, \mathbf{y}_{ij}) =$ sample containing *j* observations of Y_i $\mathbf{w}_i = (\mathbf{w}_{i1}, \mathbf{w}_{i2}, \dots, \mathbf{w}_{ij}) =$ sample containing *j* observations of W_i

 $\mathbf{z}_i = (\mathbf{z}_{i1}, \mathbf{z}_{i2}, \dots, \mathbf{z}_{ij}) =$ sample containing *j* observations of Z_i

3.3. Symbols

 μ = population mean

 σ = population standard deviation

 σ^2 = population variance

 ρ = population correlation coefficient

n = sample size

 $N(\mu, \sigma^2)$ = univariate normal distribution with the mean μ and the variance σ^2 .

 $N_2(\mu, \Sigma)$ = bivariate normal distribution with the mean vector μ and the covariance matrix Σ .

3.4. Selection of Models and Parameter Values

In thiamine-deficient young herring gulls, a typical endogenous activity of α -ketoglutarate was N(26, 4²), a typical maximum activity was N(35, 4²), a typical correlation between the endogenous and the maximum activity was 0.8, and a typical sample size was 10 [3]. Unfortunately, the standard deviation for a wider range of mean endogenous activities could not be determined from the available herring gull data. There are, however, at least three possible alternatives:

- 1) the standard deviation is constant for all mean endogenous activities,
- the standard deviation is proportional to each mean endogenous activity (*i.e.* the coefficient of variation is constant for all mean endogenous activities),
- 3) an intermediate of the two first alternatives.

Here, alternatives (1) and (2) were simulated in two respective models (model 1 and model 2) by random sampling from a bivariate normal distribution, where (X_i, Y_i) were $N_2(\mu_i, \Sigma_i)$, i = 1, 2. The models were assumed to be valid for mean endogenous activities between 13 and 29. Outside this interval, there are some obvious violations to the models. For example, lower endogenous activities cannot have a standard deviation of 4, and higher endogenous activities are suspected to be more strongly correlated with the maximum activity. In Model 1, X_i had a constant standard deviation of 4, whereas in Model 2, the standard deviation of X_i was $4\mu_{Xi}/35$, *i.e.* proportional to the mean. In both models, Y_i had a mean of 35 and a standard deviation of 4, and ρ_{XiYi} was 0.8.

3.5. Restrictions of the Models

The differences between μ_{X1} and μ_{X2} were chosen so that powers of one (too large difference) or zero (too small difference) were avoided. The largest latency observed in field material (an adult herring gull) was 96% [3], which corre-



sponds to an endogenous activity of ~1.4. Hence, occasional x-values below 1, and the corresponding y-values, were replaced by a new x and y pair from the same distribution until a complete sample (n = 10) was obtained. By this procedure, endogenous activities were drawn from a truncated normal distribution. Under the specified conditions, the frequency of replaced x and y pairs was at most ~1‰ (for the lowest μ_{X1} in Model 1). The random variables W_i and Z_i , generated in Model 1 and Model 2, were not normally distributed, and accordingly the respective t-statistics were not t-distributed under H₀ (equal means). By visual inspection of histograms, the distributions of the *t*-statistic under H₀ were, however, indistinguishable from the corresponding *t*-distribution. Hence, the violation to the normality condition was considered negligible. In practice, for the moderate sample sizes used in many enzyme biomarker investigations, the usual case is that none of the four variables X, Y, W, or Z deviates significantly from normality, as tested e.g. with the Shapiro-Wilk normality test.

3.6. Simulation Procedure

- 1) Select a μ_{X1} value and compute μ_{X2} .
- 2) Generate the samples $(\mathbf{x}_1, \mathbf{y}_1)$ and $(\mathbf{x}_2, \mathbf{y}_2)$, with 10 observations each, by random sampling from the two bivariate normal distributions specified by the model.
- 3) Compute the corresponding samples \mathbf{w}_1 , \mathbf{w}_2 , \mathbf{z}_1 , and \mathbf{z}_2 .
- 4) Test the hypothesis $\mu_{W1} = \mu_{W2}$ with Student's *t*-test. Record the *t*-value.
- 5) Test the hypothesis $\mu_{Z1} = \mu_{Z2}$ with Student's *t*-test. Record the *t*-value.

- 6) Repeat steps 2–5 10,000 times.
- 7) Determine the proportion of latency *t*-values that are larger than $t_{18,97.5}$.
- 8) Determine the proportion of TPP-effect *t*-values that are larger than $t_{18.97.5}$.
- 9) Repeat steps 1-8 until all μ_{X1} values have been investigated.

4. RESULTS AND DISCUSSION

Each respective proportion of *t*-values larger than $t_{18,97.5}$ obtained in the simulation is an estimate of the power of the *t*-test under the given conditions. In Model 1 (constant standard deviation), latency was associated with higher power than TPP-effect (Table 1), whereas in Model 2 (constant coefficient of variation), the two measures were associated with equal power (Table 2). The power difference in Model 1 was more pronounced for large latencies and TPP-effects, *i.e.* when the endogenous activity was small compared with the maximum activity. These results illustrate the general principle that normalisation to a variable with higher relative precision.

Approximately 35–45% of the power difference in Model 1 was due to the fact that the empirical 97.5% percentile of the *t*-values of the TPP-effect under H₀ was somewhat lower than $t_{18,97.5}$ (not shown). Maybe, the TPP-effect has this property because it is a less balanced measure than the latency, whose empirical 97.5% percentile of the *t*-values under H₀ was equal to $t_{18,97.5}$ (not shown). In Model 2 the empirical 97.5% percentile of the *t*-values under H₀ was equal to $t_{18,97.5}$ for both the latency and the TPP-effect (not shown).

<i>n</i> ₁ = <i>n</i> ₂	μ_{X1} [nmol/min/mg]	μ_{X2} [nmol/min/mg]	$oldsymbol{\mu}_{W1}$ [%]	μ _{W2} [%]	μ _{Z1} [%]	μ _{Z2} [%]	Power of <i>W</i> ±95% CI ^a	Power of Z ±95% CI ^a
10	13	16	62.9	54.3	169	119	0.60±0.010	0.48±0.010
10	14	17	60.0	51.4	150	106	0.61±0.010	0.51±0.010
10	15	18	57.1	48.6	133	94.4	0.63±0.010	0.55±0.010
10	16	19	54.3	45.7	119	84.2	0.65±0.009	0.58±0.010
10	17	20	51.4	42.9	106	75.0	0.66±0.009	0.61±0.010
10	18	21	48.6	40.0	94.4	66.7	0.68±0.009	0.63±0.010
10	19	22	45.7	37.1	84.2	59.1	0.70±0.009	0.66±0.009
10	20	23	42.9	34.3	75.0	52.2	0.71±0.009	0.68±0.009
10	21	24	40.0	31.4	66.7	45.8	0.72±0.009	0.69±0.009
10	22	25	37.1	28.6	59.1	40.0	0.73±0.009	0.71±0.009
10	23	26	34.3	25.7	52.2	34.6	0.74±0.009	0.72±0.009
10	24	27	31.4	22.9	45.8	29.6	0.74±0.009	0.73±0.009
10	25	28	28.6	20.0	40.0	25.0	0.75±0.009	0.74±0.009
10	26	29	25.7	17.1	34.6	20.7	0.75±0.009	0.74±0.009

Table 1. Results of Model 1

^a The confidence intervals (CI) are based on the 10,000 simulations for each combination of μ_{X1} and μ_{X2} .

$n_1 = n_2$	μ_{X1} [nmol/min/mg]	μ_{X2} [nmol/min/mg]	μ _{W1} [%]	μ _{W2} [%]	μ _{Z1} [%]	μ _{Z2} [%]	Power of <i>W</i> ±95% CI ^a	Power of Z ±95% CI ^a
10	13	15	62.9	57.1	169	133	0.98±0.003	0.98±0.003
10	14	16	60.0	54.3	150	119	0.97±0.004	0.97±0.004
10	15	17	57.1	51.4	133	106	0.95±0.005	0.95±0.005
10	16	18	54.3	48.6	119	94.4	0.92±0.005	0.92±0.005
10	17	19	51.4	45.7	106	84.2	0.89±0.006	0.89±0.006
10	18	20	48.6	42.9	94.4	75.0	0.86±0.007	0.86±0.007
10	19	21	45.7	40.0	84.2	66.7	$0.82{\pm}0.008$	0.81±0.008
10	20	22	42.9	37.1	75.0	59.1	0.78±0.008	0.78±0.008
10	21	23	40.0	34.3	66.7	52.2	0.75±0.009	0.74±0.009
10	22	24	37.1	31.4	59.1	45.8	0.71±0.009	0.71±0.009
10	23	25	34.3	28.6	52.2	40.0	0.67±0.009	0.68±0.009
10	24	26	31.4	25.7	45.8	34.6	0.64±0.010	0.64±0.010
10	25	27	28.6	22.9	40.0	29.6	0.60±0.010	0.60±0.010
10	26	28	25.7	20.0	34.6	25.0	0.57±0.010	0.57±0.010
10	27	29	22.9	17.1	29.6	20.7	0.54±0.010	0.54±0.010

Table 2Results of Model 2.

^a The confidence intervals (CI) are based on the 10,000 simulations for each combination of μ_{XI} and μ_{X2} .

No conditions were found, under which TPP-effect was associated with higher power than latency. Other models may, however, still be defined and investigated.

CONCLUSION

The results of two models show that proportion of enzyme without cofactor is associated with equal or higher statistical power than relative increase in enzyme with cofactor. The power difference can be as high as 0.12. This result should be of interest to researchers measuring the degree of missing cofactor in cofactor-dependent enzymes, since other ways of increasing statistical power by the same amount often are expensive.

CONFLICT OF INTEREST

The authors confirm that this article content has no conflicts of interest.

ACKNOWLEDGEMENTS

I am grateful to Jan-Olov Persson, member of the Statistical Research Group at the Department of Mathematics at Stockholm University, for constructive statistical advice. The work was supported by Signhild Engkvists Stiftelse och Stiftelsen Olle Engkvist Byggmästare.

APPENDIX A:

Comparison of the Coefficients of Variation (CV) of the Latency and the TPP-effect

This is a general proof, independent of the distribution of the random variables. The only assumption is that the *delta* *method* (also known as the *method of propagation of errors*) is applicable. The properties of the random variables, described in the main text, are formalised here. For simplicity, the factor of 100, used to express the latency and TPP-effect as percent, is omitted.

Random Variables

- X = endogenous activity of a specimen
- Y = maximum activity of a specimen
- W = latency of a specimen
- Z = TPP-effect of a specimen

Symbols

- μ = population mean
- σ = population standard deviation
- σ^2 = population variance

Definitions $0 < x \le y$

$$W = \frac{Y - X}{Y}$$
$$Z = \frac{Y - X}{X}$$
$$CV(W) = \frac{\sigma_W}{\mu_W}$$
$$CV(Z) = \frac{\sigma_Z}{\mu_Z}$$

$$z \ge 0$$
$$W = \frac{Z}{Z+1}$$

Investigation of the Coefficients of Variation of the Latency and the TPP-effect

Using the delta method, we obtain:

$$\mu_W \approx \frac{\mu_Z}{\mu_Z + 1}$$

$$\sigma_W^2 \approx \sigma_Z^2 \cdot \frac{1}{(\mu_Z + 1)^4}$$

$$\sigma_W \approx \sigma_Z \cdot \frac{1}{(\mu_Z + 1)^2}$$

CV(W) may be expressed as a function of CV(Z) and μ_Z :

$$\operatorname{CV}(W) = \frac{\sigma_W}{\mu_W} \approx \frac{\sigma_Z \cdot \frac{1}{(\mu_Z + 1)^2}}{\frac{\mu_Z}{\mu_Z + 1}} = \frac{\sigma_Z}{\mu_Z(\mu_Z + 1)} = \operatorname{CV}(Z) \cdot \frac{1}{(\mu_Z + 1)}$$

Since $z \ge 0$, it holds that $0 < \frac{1}{(\mu_z + 1)} \le 1$.

Hence, we obtain the approximate inequality $CV(W) \leq CV(Z)$.

Conclusion

Approximately it holds that $CV(W) \leq CV(Z)$.

APPENDIX B:

Program Code in Intercooled Stata 9.2

Model 1

```
clear
quietly set more off
quietly set obs 10000
quietly gen t w=.
quietly gen t z=.
quietly set seed 573957645
forvalue j=13(1)26 {
quietly matrix m1=(j',35)
quietly matrix m2=('j'+3,35)
quietly matrix sd=(4,4)
quietly matrix corr=(1,0.8,0.8,1)
forvalue i=1(1)10000 {
quietly drawnorm x1 y1, n(10000) means(m1) sds(sd)
corr(corr)
quietly replace x_1 = . if x_1 < 1
quietly replace y1=. if x1==.
```

quietly gen obs1= n if x1 \sim =. quietly egen nonmiss1=rank(obs1) quietly replace $x_{1=}$ if nonmiss $1 \ge 11$ quietly replace y1=. if nonmiss1>=11 quietly drawnorm x2 y2, n(10000) means(m2) sds(sd) corr(corr) quietly replace $x^{2=}$. if $x^{2<1}$ quietly replace y2=. if x2==. quietly gen obs2=_n if x2~=. quietly egen nonmiss2=rank(obs2) quietly replace $x^{2=}$. if nonmiss $^{2>=11}$ quietly replace y2=. if nonmiss2>=11 quietly gen w1=100*(y1-x1)/y1 quietly gen w2=100*(y2-x2)/y2 quietly gen z1=100*(y1-x1)/x1 quietly gen z2=100*(y2-x2)/x2 quietly ttest w1==w2, unpaired quietly replace t w=r(t) in `i' quietly ttest z1==z2, unpaired quietly replace t z=r(t) in 'i' quietly drop x1 y1 w1 z1 obs1 nonmiss1 x2 y2 w2 z2 obs2 nonmiss2

}

disp "x_mean 1=" `j' disp "x mean 2=" `j'+3 quietly count if t w>invttail(18,0.025) disp "w power=" r(N)/10000quietly cii 10000 r(N), exact disp "95% CI: " r(lb) " " r(ub) quietly count if t z>invttail(18,0.025) disp "z power=" r(N)/10000quietly cii 10000 r(N), exact disp "95% CI: " r(lb) " " r(ub) disp "" } Model 2 clear quietly set more off quietly set obs 10000 quietly gen t w=. quietly gen t_z=. quietly set seed 573957645 forvalue j=13(1)27 {

Proportion Versus Relative Increase

```
quietly matrix m1=(j',35)
quietly matrix m2=(j+2,35)
quietly matrix sd1=(4*(j')/35,4)
quietly matrix sd2 = (4*(j+2)/35,4)
quietly matrix corr=(1,0.8,0.8,1)
forvalue i=1(1)10000 {
quietly drawnorm x1 y1, n(10000) means(m1) sds(sd1)
corr(corr)
quietly replace x_1 = . if x_1 < 1
quietly replace y1=. if x1==.
quietly gen obs1= n if x1\sim=.
quietly egen nonmiss1=rank(obs1)
quietly replace x_{1=} if nonmiss1 \ge 11
quietly replace y1=. if nonmiss1>=11
quietly drawnorm x2 y2, n(10000) means(m2) sds(sd2)
corr(corr)
quietly replace x^{2=}. if x^{2<1}
quietly replace y2=. if x2==.
quietly gen obs2= n if x2 \sim =.
quietly egen nonmiss2=rank(obs2)
quietly replace x^{2=}. if nonmiss^{2>=11}
quietly replace y2=. if nonmiss2>=11
quietly gen w1=100*(y1-x1)/y1
quietly gen w2=100*(y2-x2)/y2
quietly gen z1=100*(y1-x1)/x1
quietly gen z2=100*(y2-x2)/x2
quietly ttest w1==w2, unpaired
```

quietly replace t w=r(t) in 'i' quietly ttest z1==z2, unpaired quietly replace t z=r(t) in `i' quietly drop x1 y1 w1 z1 obs1 nonmiss1 x2 y2 w2 z2 obs2 nonmiss2 } disp "x mean 1=" `j' disp "x mean 2=" j'+2quietly count if t_w>invttail(18,0.025) disp "w power=" r(N)/10000 quietly cii 10000 r(N), exact disp "95% CI: " r(lb) " " r(ub) quietly count if t z > invttail(18, 0.025)disp "z power=" r(N)/10000quietly cii 10000 r(N), exact disp "95% CI: " r(lb) " " r(ub) disp "" }

REFERENCES

- Karapinar T, Dabak M, Kizil O. Thiamine status of feedlot cattle fed a high-concentrate diet. Can Vet J 2010; 51: 1251-3.
- [2] Blair PV, Kobayashi R, Edwards HM, Shay NF, Baker DH, Harris RA. Dietary thiamin level influences levels of its diphosphate form and thiamin-dependent enzymic activities of rat liver. J Nutr 1999; 129: 641-8.
- [3] Balk L, Hägerroth P-Å, Åkerman G, et al. Wild birds of declining European species are dying from a thiamine deficiency syndrome. Proc Natl Acad Sci USA 2009; 106: 12001-6.

Received: June 12, 2012

Revised: August 23, 2012

Accepted: August 28, 2012

© Tomas Hansson; Licensee Bentham Open.

This is an open access article licensed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/3.0/) which permits unrestricted, non-commercial use, distribution and reproduction in any medium, provided the work is properly cited.