How Much would you Rely on Statistical Data? The Power of Statistical Analysis towards Truth as Well as Lie

Petridis AK*

Department of Neurosurgery, Sana Kliniken, Zu den Rehwiesen 9, P.O Box: 47055, Duisburg, Germany

Abstract

Background: Statistics is a powerfull tool in the hands of an expert but in the hands of a fool it can result in devastating conclusions, since it can fool the whole scientific world.

Method: A simulation of 340 patients with growth hormone overexpression and visual disturbance is compared with a control group of normal expressing growth hormone in healthy individuals. Different statistical analyses have been performed.

Results: All statistical analyses performed here, show a highly significant correlation of growth hormone overexpression and visual disturbance.

Conclusion: Statistically it seems that growth hormone is a cause of visual disturbance. We know that growth hormone by itself is not causing blindness but the disease hidden behind this overexpression, namely macro adenomas does. Would we not be aware of the causative disease we would easily come to the conclusion that growth hormone is strongly associated with visual disturbance. If the knowledge about a pathologic entity is incomplete a statistical analysis alone can mislead the scientific world to very strange conclusions.

Keywords: Confidence interval; Statistical analysis, Statistical testing

Introduction

No observational retrospective or prospective study can convince an investigator without a neat and clean statistical analysis. However, statistics apart of being the investigators best friend can also become a vicious enemy when not used appropriately. It is like a weapon with enormous power in the hands of a child. You never know what the outcome will be. Because of the power statistics have and their representation of truth we rely almost blindly on our analyses. In the present paper it is shown how convincing statistical results can be no matter how far away the outcome can be from the truth.

Methods

The percentage of patients with visual disturbance and growth hormone positive macroadenomas is about 85%. The fact that the group of patients with growth hormone overexpression is patients with macroadenomas will be neglected. The diagnosis macroadenoma will be an unknown factor in the statistical analysis. The percentage of individuals with visual disturbance in the normal population is about 10%. In this simulation a fictive number of two collectives, one with patients with growth hormone overexpression (IGF-1) (N=340) and one without growth hormone overexpression in a healthy population (1022), and the proportion of visual disturbance in both collectives is calculated. A retrospective analysis of a possible correlation of growth hormone overexpression (IGF-1) and visual disturbance has been performed.

Results

Inference for a single proportion in patients with IGF-1 overexpression and visual disturbance compared to a control group of individuals without IGF-1 overexpression

It is observed that 85% of patients with IGF-1 overexpression (N=340), suffer from visual disturbance. In the control collective of 1022 patients admitted to a hospital with diagnoses different than IGF-1 overexpression (i.e. trauma with isolated fractures of extremities and disc hernias of the spine) visual disturbances could be seen in about 10% of individuals. Is the difference in the percentage of visual disturbances in the IGF-1 overexpression group statistically significant compared to the control?

First of all an inference test for a single proportion will be used to evaluate a significant correlation in IGF-1 overexpression and visual disturbance.

Before statistical anaylsis is performed we have to evaluate if the conditions for our sampling distribution p is being nearly normal (Table 1).

1. Independence: The sample observation is assumed to be independent from each other.

2. There is expected to see at least 10 successes and 10 failures on each collective.

Successes in the IGF-1 overexpression group=np=340 × 0.85=289 >10

	Visual Disturbance	Normal Vision	Total
IGF-1 (+)	290	50	340
IGFnorm	100	922	1022
	390	972	1362

Table 1: IGF-1 overexpression (IGF-1 (+)) and visual visual disturbance. IGF norm: physiological levels of IGF.

*Corresponding author: Dr. Athanasios K. Petridis, Department of Neurosurgery, Sana Kliniken, Zu den Rehwiesen 9, 47055, Duisburg, Germany, Tel: 492037332425; Fax: 492037330; E-mail: opticdisc@aol.com

Failures in the IGF-1 overexpression group=n (1-p)=340 × 0.15=51 >10 and

Successes in the IGF-1 normal group=np=1022 × 0.09=92 >10

Failures in the IGF-1 overexpression group= n (1-p)=1022 × 0.91=930 >10

The success-failure condition is met too therefore the sampling distribution of IGF-1 overexpression is nearly normal with a mean of p (0.85) and a standard error of SE= $\sqrt{}$ (p (1-p))/$\sqrt{n}$= $\sqrt{}$ (0.85 × 0.15)/340=0.019.

The 95% confidence interval (CI) is: Point estimate (p) ± z* × SE=0.85 ± 1.96 × 0.019=0.85 ± 0.037.

The confidence interval for the IGF-1 overexpression group with an 85% correlation with visual disturbance is (0.813, 0.887).

The mean p° of the control (which is also normally distributed) is 0.09 and falls far away from the 95% CI of the IGF-1 overexpression group indicating a statistically highly significant difference in the occurrence of IGF-1 overexpression and visual disturbance compared to the control group where IGF-1 is normaly expressed.

The CI for the control group is: 95%. CI controls=point estimate (p°) ± 1.96 × SE= 0.09 ± 1.96 × ($\sqrt{}$(0.09 × 0.91)/1022)== 0.09 ± 1.96 × 0.00028=0.09 ± 0.0054. The confidence interval at 95% in the control group is: (0.0846, 0.09054)

Another statistical approach is hypothesis testing for our proportion of 85% (IGF-1(+) and visual disturbance).

The Z score is: Z=point estimate (p)–null value (p°)/SE=0.85– 0.09/0.019=39 which is again a highly significant value for a p value of <<< 0.001.

Again the correlation of IGF-1 overexpression and visual disturbance is highly significant compared to control group of patients.

And again another statistical test with testing the hypothesis Ho: p1=p2. Sample distribution of the difference of two proportions.

In this case the null hypothesis is that the proportion of individuals with visual disturbance in the IGF-1 overexpressing group (p1) is the same as in the control group (p2). So, Ho: p1–p2=0

The alternative hypothesis is p1–p2 ≠ 0 or in our case p1–p2 > 0. This means our alternative hypothesis states that visually disturbed individuals are likely to overexpress IGF-1.

The standard error SE is SE=$\sqrt{}$((pIGF(+)(1-pIGF(+)))/(n(visual dis))+(pIGF(-)(1-p(IGF(-))/(n (normal seight)))

There is just one problem. In the null hypothesis the proportion of visual disturbance would be the same in both groups. In the standard error equation the p (IGF1 (+) and the p(IGF1norm) would be the same. So the exposure rate (the rate of IGF-1 overexpression), assuming the null hypothesis is true, should be a pooled estimate p which is calculated by pooling the results of both samples.

Pooled estimate of a proportion when null hypothesis Ho: p1=p2 is true is:

p(pooled)=(number of successes)/(number of cases)=290 + 100/340 + 1022=390/1362=0.28

The p(pooled) is used to compute the standard error.

The standard error SE is SE=$\sqrt{}$((p(pooled)(1-p(pooled))/(n(IGF-1

overexpress))+(p(pooled)(1-p(pooled))/(n (nomal IGF-1))))=(0.28 × 0.72)/340 + (0.28 × 0.72)/1022=0.028

And the z score is: Z= ((p (IGF-1(+)–p(IGF-1norm))–null value)/ SE=(0.85- 0.09-0)/0.028=27.14

This z score correlates to a p value <<< 0.001 and therefore the null hypothesis can be rejected.

Discussion

In each different statistical approach which was done the null hypothesis, that IGF-1 overexpression is not correlated with visual disturbance could be rejected. The alternative hypothesis that IGF-1 expression leads to visual disturbance seems to be right. Statistically IGF-1 overexpression leads surely to visual disturbance. Antagonizing growth hormone would even lead to a regress of visual disturbance. Is IGF-1 a cause for blindness? Or is IGF1 overexpression by itself disturbing vision? This question cannot be answered by our statistical analysis because the study is just observational and not experimental. If it would not be IGF-1 overexpression but a completely new molecule, which we would try to correlate with visual disturbance, the chance to convince the audience that this new molecule causes blindness would be high, because of our statistical analysis and the great z scores. The small but essential detail that the study design by itself is weak would probably not even be realized.

Observational studies have to be based on logical assumptions which on their turn are basing on knowledge we already gathered through a number of studies otherwise they can mislead us in a direction completely opposite to the truth. Statistics or numbers do not lie, but they cannot be blindly relied on. A good statistical analysis on a bad study design can lead to the worst possible result which is getting the study published and mislead a chain of scientist or clinicians to a dark area. It is possible to come to wrong assumptions as an author as well. Because the author of a manuscript relies also on his statistical numbers he by himself can be convinced that his observation is true, can stand alone and does not need further evaluation.

The aim of this short paper is not to ban observational studies. The aim is to make clear that an observation alone, based on neat statistical data is not worth much unless the scientist knows already where the path starts and therefore where it probably leads. The best study design for completely new scientific landscapes is always the experimental one the prospective study. The retrospective study on its turn fits much more in evaluating details of something, which is already based on stable scientific foundations.

Statistics and numbers alone are representing a mathematical truth but are dangerous if blindly relying on them. After all the statistical significant analysis of jumping out of a plane without a parachute cannot be defined easily if at all. This should be a great example that even when statistically there is no increased risk of death by skydiving without a parachute no one would rely on that numbers [1]!

References

1. Diez DM, Barr CD, Cetinkaya-Rundel M (2012) OpenIntro Statistics. Openintro. org, (2ndedn) 268-290.