

## Possible statistical problems of the original paper titled “Predictors of difficult intubation”: a bad start

In science, statistics is very important in the original article.

It is not only proves the author's conclusions, but also gives accreditation to the paper.

Additionally, statistics, which had a basis in mathematics, is essential in modern sciences. Of course it provides a concrete basis for an original, scientific article in medicine.

Basically, there are two kinds of tests in order to testify the null hypothesis ( $H_0$ ) in medical statistics. One is a test to show the difference, whereas the other is to prove the equivalence – test for no difference – between control and experimental group.

For the difference test, a null hypothesis ( $H_0$ ) should be made. Null hypothesis is usually asserts that "there are no differences between two groups. Nevertheless if there is a difference between two groups, it is originated from the sampling variance" [1].

To guarantee no differences between two sampled groups, variance from sampling should be minimized. It can be realized by random sampling in a clinical study. If we increase the sample size by random sampling, there would be a lessening of differences in demographic data between groups. Of course this means that there should be no differences in the factors that could affect the results of the experiment between the two groups [1].

The next step is calculating a probability (P value) of achieving no difference in the result between groups. If it is below 5% ( $P < 0.05$ ), it means that no difference between groups, if ever, would take place very seldom. Hence, it is reasonable to reject the null hypothesis. Then we can say that there is a significant difference between groups statistically [1].

It is time to review the paper “Predictors of difficult intubation defined by the intubation difficulty scale (IDS): predictive value of 7 airway assessment factors.” written by Seo et al. [2]. The aim of the paper is to find the preoperative airway risk factors that can predict the difficult intubation more efficiently.

The two groups are divided by the criteria of intubation

difficulty score, which was measured by the experimenter, who had administered endotracheal intubation. The demographic differences between the two groups are shown in Table 3 of the aforementioned paper. There are already three statistically different factors (age, intubation duration, lowest SaO<sub>2</sub> level) between the two groups. From this information, what kind of a null hypothesis could be built? To make a null hypothesis, there should be no differences in the demographic data between the two groups. If there are different factors between the two groups with regards to the demographic data, it is very hard to rule out their influence on the result. Additionally there is a significant relationship between the criteria that divides a population into two groups and the results itself, which is shown in this paper. There are many clinical tests to predict a possible airway risk. Of course, these types of test are officially accepted as being useful in predicting a possible difficulty in endotracheal intubation. Most of them appear in this paper. Considering the link between the criteria by which the two groups are divided and the experimental results, it is easy to infer that there are big differences in the results between the two groups.

Finally, Table 5 shows that the total airway score is the most useful factor in predicting the difficulty intubation, followed by the upper lip biting test. However most tests have false positive and negative traits. Therefore, it is very hard to say that one test is better than another just for the reason of high sensitivity. Additionally, the author should show why he used the method of multiple logistic regression analysis in Table 5. Odds ratios are hardly used in testifying a usefulness of a clinical test.

### Sae-Cheol Oh

Department of Anesthesiology and Pain Medicine, College of Medicine, The Catholic University of Korea, Daejeon St. Mary's Hospital, Daejeon, Korea

E-mail: [oscaen0822@gmail.com](mailto:oscaen0822@gmail.com)

---

© This is an open-access article distributed 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 original work is properly cited.

## References

1. Kirkwood BR, Sterne JA. *Essential Medical Statistics*. 2nd ed. Oxford, Blackwell publishing. 2003, pp 58-86.
2. Seo SH, Lee JG, Yu SB, Kim DS, Ryu SJ, Kim KH. Predictors of difficult intubation defined by the intubation difficulty scale (IDS): predictive value of 7 airway assessment factors. *Korean J Anesthesiol* 2012; 63: 491-7.

## In Response

To predict a difficult intubation (DI) before anesthetic induction is a one of the anesthesiologists wishes (in order to predict a DI, we compared characters of difficult intubated patients with normally intubated patients using the total airway score [TAS] and airway factors).

Our study was a retrospective, cross-section analysis as well as an observational study in which patients were classified into IDS > 5 group and IDS < 5 group according to 7-criteria IDS as the diagnostic criteria for difficult intubation. Our study was not a randomized trial test.

Dr. Oh pointed to three statistically different factors (age, intubation duration, lowest SaO<sub>2</sub> level) between the two groups (Table 3) and there should be no difference in the demographic data between the two groups in order to make a null hypothesis. I do not think he completely understands our paper. Our groups were classified by the 7-criteria of IDS. Thus, there should be a difference between the two groups (Table 3). The intubation duration and the lowest SaO<sub>2</sub> level criteria are 2 of 7-criteria IDS. Characteristic data (Table 3) [1] of the two groups were not randomly sampled. Therefore, we do not think that the null hypothesis should be applied to Table 3. We did not mention anything about the age difference in this article; however, the age difference between the two groups can be a one of the variables to predict DI. In this article, we focused on the study regarding the TAS and airway factors predictability of DI. We must consider calibrating the odd's ratio of airway factors

adding age variables data in the next study,

Categorical data of TAS and airway factors were analyzed using the Chi Squared test (Table 3 and Table 4) [1].

Clinical research on the prediction of a DI should focus on a multivariable analysis to predict DI. Ranking the predictors was based on the odd's ratio in other papers [2]. (Multivariable logistic regression was used to measure the relationship between TAS and airway factors, and IDS > 5 group (Table 5) [1].

I agree with him that multiple abnormal tests predicting DI are better than any single test. (Combinations of tests may lead to better anticipation of DI in comparison to each test alone.)

We suggested that total airway score (TAS) > 6 was best to predict DI; however, if the circumstance does not allow using TAS, a lip bite test is useful than any single test.

Suk-Hwan Seo, Jeong-Gil Lee, Soo-Bong Yu, Doo-Sik Kim, Sie-Jeong Ryu, and Kyung-Han Kim

Department of Anesthesiology and Pain Medicine, Kosin University College of Medicine, Busan, Korea

### Kyung-Han Kim

Department of Anesthesiology and Pain Medicine, Kosin University College of Medicine, Busan, Korea  
E-mail: kimkh@ns.kosinmed.or.kr

## References

1. Seo SH, Lee JG, Yu SB, Kim DS, Ryu SJ, Kim KH. Predictors of difficult intubation defined by the intubation difficulty scale (IDS): predictive value of 7 airway assessment factors. *Korean J Anesthesiol* 2012; 63: 491-7.
2. Shiga T, Wajima Z, Inoue T, Sakamoto A. Predicting difficult intubation in apparently normal patients: a meta-analysis of bedside screening test performance. *Anesthesiology* 2005; 103: 429-37.