Doctors' Statistical Literacy: A Survey at Srinagarind Hospital, Khon Kaen University†

MALINEE LAOPAIBOON, M.Sc.*,
PISAKE LUMBIGANON, M.D.**,
STEPHEN D WALTER, M.D.***

Abstract

Medical doctors need to keep abreast of new developments in medicine. This is often done by reading medical journals and carrying out research activities that require an understanding of statistical methods. This study was designed to assess the knowledge of statistics among doctors in Thailand. A pretested, self-administered questionnaire with nine multiple-choice questions on basic statistical issues was used. In a survey of university hospital staff, there were 365 doctors, including 156 specialists, 152 residents and 57 final year medical students (externs). The overall response rate was 40.0 per cent. The overall median number of correct answers was 4.0 (95% CI 3.0, 4.0). Specialists had a significantly higher median score, 4.0 than residents, and externs, 3.0's, ($p = 0.02$). Respondents who had previously attended statistical workshops had a significantly higher median score (5.0) than those who had not (3.0) ($p < 0.01$). These results indicate that doctors in our hospital have insufficient knowledge of the basic statistical concepts that are commonly used in medical journals. Continuing education in statistics for doctors during residency and post doctoral training must be given serious consideration.

To be a competent professional in clinical work, doctors need to keep abreast of new developments in medicine, usually by reading medical journals. They should also be able to critically evaluate the results of original research. Medical research is one of the main foundations of medical knowledge, influencing diagnosis and treatment. Research that is badly designed or misleadingly analyzed may lead to wrong diagnostic or therapeutic decisions and so put patient health or even lives

* Department of Biostatistics and Demography, Faculty of Public Health,
** Department of Obstetrics and Gynecology, Faculty of Medicine, Khon Kaen University, Khon Kaen 40002, Thailand.
*** Department of Clinical Epidemiology and Biostatistics, McMaster University Health Sciences Centre, 1200 Main Street West, Hamilton, Ontario, L8N 3Z5 Canada.

† This study was presented at the Inclen XI Annual Meeting, Cairo, Egypt, 24-29 January, 1993 and International Epidemiological Association 13th Scientific Meeting, Sydney, Australia, 26-29 September, 1993.

at risk[1]. Moreover, a high proportion of papers in the medical journals involve some form of statistical analysis. Several researchers[2-8] have reviewed the statistical techniques used in medical publications, and they have shown frequent errors and misuse of statistics. Misuse of statistical techniques may lead to incorrect conclusions, which may then be widely and uncritically reported, thus adversely affecting subsequent research and patient care. Furthermore, some doctors also carry out researches and are responsible for statistical analyses themselves. Thus, it is clear that doctors need an understanding of fundamental statistical methods.

Fundamental statistical concepts, such as standard deviation (SD), standard error (SE), p-value, confidence interval (CI), correlation coefficient (r), indices of performance for diagnostic test (such as sensitivity and positive predictive value) and relative risk, are commonly presented in medical journals. A number of studies on the statistical knowledge of doctors have been conducted in some developed countries. They have all suggested that doctors' statistical knowledge of basic methods and concepts is poor[7-13]. There is, however, no information about the level of statistical knowledge of doctors in developing countries.

In Thailand, members of the medical profession, especially in medical schools, are becoming increasingly interested in reading medical papers and doing research. Similar to most other developing countries and developed countries, Thai medical schools include an introductory biostatistics course in the preclinical years. Doctors usually need the statistical concepts for appraising medical papers and doing research during their clinical rotation and residency training, but by that time, they tend to have forgotten these concepts. To reorganize the medical curriculum and arrange continuing education in biostatistics during residency training, we require information on this issue for Thai doctors. The present study was designed to assess the knowledge level of doctors at Srinagarind Hospital, a teaching hospital of Khon Kaen University, Thailand. The emphasis was on elementary statistical concepts frequently encountered in medical journals.

PARTICIPANTS AND METHOD

The study population consisted of all the 365 doctors working at Srinagarind Hospital, including 156 specialists, 152 residents and 57 externs (final year medical students) from February to June 1992. A questionnaire was developed in the Thai language with a goal of testing the subject's knowledge of some statistical concepts frequently encountered in medical journals. The questionnaire had two parts. The first part asked three questions related to the amount of formal course-work in statistics the subjects had received during their undergraduate and postgraduate medical training, and the subject's perception of statistical roles in their clinical practice. The second part asked nine multiple-choice questions aimed at assessing knowledge of several basic statistical concepts (presented in Appendix). Of these questions, three were modified from the 1983 survey of Danish physicians[13], one question was modified from the survey of physicians at Harvard Medical School teaching hospitals[14], the other five questions were developed by us. The questions addressed the meanings of standard deviation (SD), standard error (SE), correlation coefficient (r), sensitivity and positive predictive value of a diagnostic test, the interpretation of a p-value, relative risk for a risk factor and a protective factor, and the calculation of positive predictive value. The subject received one mark for each correct answer.

Three biostatisticians from the Department of Biostatistics and Demography in the faculty of Public Health and one faculty, trained in clinical epidemiology, from the faculty of Medicine in Khon Kaen University reviewed the statistical questions for clarity and appropriateness. After refining the questionnaire according to their suggestions, a pretest was conducted to assure its clarity. A final version of the questionnaire, with a covering letter explaining the purpose of the study was distributed to each subject in March, 1992. We expected to receive the anonymous replies within one month. As the response rate was so low, we had to extend a deadline of the replies to four months (June, 1992). Because of the anonymous replies, we could not know the characteristics of the nonrespondents.

By using the Lilliefors test[15], we found that overall number of correct answers was not normally distributed, but had positive skewness (Fig. 1). Therefore, medians and 95 per cent confidence intervals (CI) for the medians[16] were used to describe the level of statistical knowledge. The Median test was employed to assess the statistical significance of the differences between two or more median differences at a level of 5 per cent. After analyzing the data, we distributed pamphlets

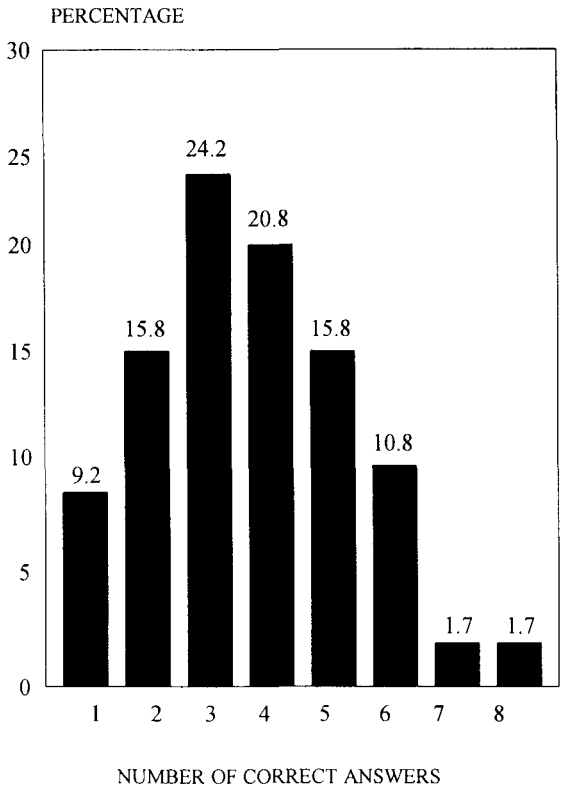

Fig. 1. Percentage distribution of respondents by number of correct answers.

that discussed in detail the correct and incorrect answers for each question, with a letter of appreciation to each respondent.

RESULTS

An overall response rate of 40 per cent (146 out of 365 subjects) was achieved, with different response rates among three groups of subjects (Table 1). The percentage of correct responses to individual questions ranged from 18.6 in the question concerning p-value to 88.3 in the question concerning sensitivity (Table 2). The median number of correct answers from the set of nine questions was 4.0 (95% CI 3.0, 4.0). Two respondents (1.7%) scored 8 points, the highest value attained. Seventy per cent of all respondents had scores less than 5.0 (Fig. 1). Table 3 (a) shows the specialists had a median score of 4.0 which was statistically significant higher than the median score of 3.0 for the residents and the externs (p = 0.02).

Table 1. Percentages of respondents by level of professional practice.

Level of professional practice	Per cent (No. of response/total subjects)
Specialist	34.0 (53/156)
Resident	38.2 (58/152)
Extern*	61.4 (35/57)
Total	40.0 (146/365)

* Extern = final year medical student

Table 2. Percentage of correct answers of respondents by statistical concept.

Statistical concept (Question number)	Per cent correct
- interpretation of p value(3)	18.6
- calculation of positive predictive value (PPV)(7)	19.7
- meaning of SE(2)	22.8
- meaning of correlation coefficient(4)	27.1
- interpretation of relative risk for a risk factor(8)	35.7
- meaning of SD(1)	37.2
- interpretation of relative risk for a protective factor(9)	46.5
- meaning of PPV(6)	69.9
- meaning of sensitivity(5)	88.3

Table 3. Median and 95 per cent CIs for median number of correct answers.

	Median	95% CI for median
a). by professional level		
Specialist	4.0	3.0, 5.0
Resident	3.0	2.0, 4.0
Extern	3.0	3.0, 4.0
b). by prior attendance at statistical workshop		
Ever	5.0	4.0, 6.0
Never	3.0	3.0, 4.0

All respondents stated that they had taken an introductory course in statistics, mostly in the first undergraduate year. However, only 24 per cent reported prior attendance at statistical workshops during their clinical practice; this group had a median score of 5.0, which was statistically significant higher than the median score of 3.0 for the respondents who had not attended a workshops (p

< 0.01). The 95 per cent CIs for the median scores of the two groups are also presented in Table 3 (b). Ninety-six per cent of respondents stated that they needed statistical knowledge for their work. Of these, 92 per cent further reported that their statistical knowledge was not enough to interpret the results of statistical analysis found in medical journals.

DISCUSSION

Similar to other studies[7-13], our findings suggest that most doctors have inadequate knowledge of elementary statistics to correctly interpret the results of statistical analyses frequently found in medical journals.

Less than 20 per cent of the respondents correctly answered the question concerning the interpretation of p-values, a concept that is frequently encountered in medical journals. Our findings accord well with those of Wulff et al[13] who found a corresponding figure of only 13 per cent among Danish doctors. This problem is very important, because most doctors often use p-values to summarize their study results but incorrectly understand their meaning. In an attempt to avoid the misuse of p-values, many journals are now requiring that results be expressed in terms of confidence intervals (CI), rather than p-values[17].

Standard deviations and standard errors are also used in many medical papers but our survey revealed that most doctors are unable to interpret these concepts. Our results also support the finding of Brown[10], who showed that the authors of medical papers seemed to know as little about

Appendix: Statistical questions

		% Answers
1.	In a study of 100 diabetics, they were characterized as "age 35 years ± 5 years (mean ± standard deviation)". Which of the following statements is the most correct answer? (N = 146)	
a.	The true mean age of diabetics lies within the interval 25-45 years with 95 per cent confidence.	17.9
b.	Most of the diabetics were aged 35 years, the remainder were aged between 30 years and 40 years.	37.9
c.	Approximately 95 per cent of the diabetics were aged between 25 years and 45 years.	37.2
d.	I do not understand the expression.	7.0
	Answer c.	
2.	A study of 200 patients with heart disease showed that "the systolic blood pressure was 120 mmHg ± 5 mmHg (mean ± standard error)". Which of the following statements is the most correct answer? (N = 146)	
a.	Approximately 95 per cent of the patients had systolic blood pressure between 115 and 125 mmHg.	49.0
b.	Mean systolic blood pressure of the patients was between 110 and 130 mmHg.	15.9
c.	We are 95 per cent confident that the true mean systolic blood pressure of the patients with heart disease lies somewhere within the interval 110 to 130 mmHg.	22.8
d.	I do not understand the expression.	12.3
	Answer c.	
3.	A controlled trial of a new treatment led to the conclusion that "it is significantly better than placebo : p value < 0.05". Which of the following statements do you prefer? (N = 146)	
a.	It has been proved that the treatment is better than placebo.	24.1
b.	If the treatment is not effective, there is less than a 5 per cent chance of obtaining such results.	18.6
c.	The observed effect of the treatment is so large that there is less than 5 per cent chance that the treatment is not better than placebo.	47.6
d.	I do not really know what a p value is and do not understand what statistical significance means.	9.7
	Answer b.	
4.	Which correlation coefficient below shows the strongest relationship between two variables? (N = 133)	
a.	+ 0.50	21.8
b.	+ 0.85	20.3
c.	+ 1.25	30.8
d.	- 0.95	27.1
	Answer d.	

		% Answers
5.	Which of the following statements is the meaning of sensitivity? (N = 145)	
a.	The proportion of persons with a positive test who are disease - free.	4.1
b.	The proportion of persons with a negative test who are disease - free.	4.8
c.	The proportion of persons with the disease who have a positive test result.	88.3
d.	The proportion of persons with a negative test who have disease.	2.8
	Answer c.	
6.	Which of the following statements is the meaning of positive predictive value? (N = 143)	
a.	The proportion of persons without disease having a positive test result.	4.2
b.	The proportion of persons with a positive test who actually have the disease of interest.	69.9
c.	The proportion of persons with a positive test who are actually normal	4.2
d.	The proportion of persons with the disease having a positive test result	21.7
	Answer b.	
7.	"A test with 80 per cent sensitivity is applied to detect a disease whose prevalence is 5/1,000, it has a false positive rate of 5 per cent" Which of the following statements is the most correct answer? (N = 142)	
a.	The positive predictive rate is less than 8 per cent.	19.7
b.	The false negative rate is less than 8 per cent.	15.5
c.	The specificity is less than 8 per cent.	9.9
d.	I do not understand the statements above.	54.9
	Answer a.	
8.	A study of risk factors of preterm delivery found that the smoking mothers were 1.5 times as likely to have a preterm delivery as nonsmoking mothers which is not statistically significant at 5 per cent level. Which of the following statements is the most correct answer? (N = 129)	
a.	One is 95 per cent confident that the relative risk of preterm delivery for smoking mothers lies between 0.25 and 1.15	24.8
b.	One is 95 per cent confident that the relative risk of preterm delivery for smoking mothers lies between 0.75 and 3.55	35.7
c.	One is 95 per cent confident that the relative risk of preterm delivery for smoking mothers lies between 1.21 and 4.65	15.5
d.	One is 95 per cent confident that the relative risk of preterm delivery for smoking mothers lies between -0.50 and 1.76	24.0
	Answer b.	
9.	A controlled trial of a new vaccine for preventing influenza revealed that members of the study group were 0.5 times as likely to have had an influenza as members of the control group, which is statistically significant at 5 per cent level. Which of the following statements is the most correct? (N = 127)	
a.	One is 95 per cent confident that the relative risk of influenza for vaccinated people lies between 0.28 and 7.56	11.0
b.	One is 95 per cent confident that the relative risk of influenza for vaccinated people lies between 2.30 and 5.00	29.1
c.	One is 95 per cent confident that the relative risk of influenza for vaccinated people lies between -2.15 and 0.90	13.4
d.	One is 95 per cent confident that the relative risk of influenza for vaccinated people lies between 0.23 and 0.89	46.5
	Answer d.	

Question 1, 2 were modified from Wulff et al (1987).
Question 3 was the same as that of Wulff et al (1987).
Question 7 was modified from Casscell et al (1978).
Question 4, 5, 6, 8 and 9 were developed by ourselves.

the correct interpretation of SD and SE as their readers.

Our results show that more than 70 per cent of the respondents answered the question on correlation incorrectly. The most commonly chosen response option (selected by 30.8% of respondents) gave a correlation value of + 1.25 (see Appendix, answer C to Q.4). This indicates that most doctors do not understand the concept of a correlation, whose value cannot exceed 1.0. Our finding is similar to those of Friedman and Phillips[11] who found that many paediatric residents in the United States were unable to answer a simple question concerning the magnitude of correlation coefficients. Less than 20 per cent of the residents correctly chose the numerical value of correlation that shows the strongest association between two variables.

More than half of the doctors could not correctly interpret the numerical value of relative risk in relation to the concepts of confidence interval and statistical significance. This might be because the question in our study was too difficult. To answer the question, the respondents had to understand not only the concept of relative risk but also the relationship between confidence interval and statistical significance for the relative risk. Our finding is in contrast to those of Weiss and Samet[18] who found that 97 per cent of physicians in a teaching hospital in the United States understood the concept of relative risk.

Although most of the doctors understood the definition of sensitivity and positive predictive value (PPV), the percentage of correct responses to the question concerning the calculation of PPV was very low (19.7%). This might be because the question was too complicated; indeed more than half of the respondents selected the answer "I do not understand the statement above". However, our result is similar to that of Casscells et al[14] who found that only 18 per cent of physicians at Harvard Medical School teaching hospitals were able to calculate PPV correctly.

Doctors who had previously attended a statistical workshop scored higher and so may be better prepared to assess medical journals. The workshops appear to have had some impact on the statistical knowledge of doctors in our hospital, although selection bias of the participants into the workshop might be another possible explanation. Most of them had been to the statistical workshop that is a one week course annually organized by the Clinical Epidemiology Unit in the Faculty of Medicine. The participants are the first 40 applicants from the health sciences faculties of Khon Kaen University and health agencies of the ministry of Public Health. The course covers some common descriptive statistics, confidence intervals and significance testing, inference on means and proportions, relative risk and odds ratio, and diagnostic test evaluation.

Two potential limitations of the results must be considered. First, the overall response rate was only 40.0 per cent in spite of repeated attempts to obtain better participation. We speculate that the non-respondents would have scored even lower than the respondents. Unfortunately, with anonymous replies, we could not examine the comparability of the respondents and the non-respondents. In fact, some of the non-respondents indicated that they did not know any statistical concepts and techniques, and were reluctant to be tested in this study. Some also stated that the statistical questions were so difficult that they preferred not to respond. Second, the questionnaire measured only limited aspects of statistical knowledge. However, the statistical concepts and expressions included in the questionnaire were those that are used most commonly in medical journals. Thus, we believe this study is sufficient to demonstrate the inadequate knowledge level of elementary statistics of doctors in Srinagarind Hospital. Since curricula of medical schools in Thailand and even other developing countries are very similar, we believe that the findings of this survey are generalizable to other developing countries.

In conclusion, the results of this study reaffirm that doctors have insufficient knowledge of the basic statistical concepts that are commonly used in medical journals. Continuing education in statistics for doctors during residency training and clinical rotation must be given serious consideration.

ACKNOWLEDGEMENTS

The authors wish to thank the study participants, Miss Natwadee Patoombal for collecting the questionnaires and recording the data and Miss Radda Wilailert for typing the manuscript.

(Received for publication on July 26, 1996)

REFERENCES

1. Altman DG, Bland JM. Improving doctors' understanding of statistics. JR Stat Soc Ser A 1991; 154: 223-67.
2. Emerson JD, Colditz GA. Use of statistical analysis in the New England Journal of Medicine. N Engl J Med 1983; 309: 709-13.
3. Goodman NA, Hughes AO. Statistical awareness of research workers in British anaesthesia. Br J Anaesth 1992; 68: 321-4.
4. Hayden GF. Biostatistical trends in Pediatrics : implications for the future. Pediatrics 1983; 72: 84-7.
5. Mainland D. The use and misuse of statistics in medical publications. Clin Pharmacol Ther 1960; 1: 411-22.
6. Reznick RK Dawson - Saunders E, Folse JR. A rationale for the teaching of statistics to surgical residents. Surgery 1987; 101: 611-7.
7. White SJ. Statistical errors in papers in the British Journal of Psychiatry. Br J Psychiatry 1979; 135: 336-42.
8. Gore SM, Jones IG, Rytter BC. Misuse of statistical methods : critical assessment of articles in BMJ from January to March 1976. BMJ 1977; 1: 85-7.
9. Berwick DM, Fineberg HV, Weinstein MC. When doctors meet number. Am J Med 1981; 71: 991-8.
10. Brown GW. Standard deviation, standard error : Which standard should we use? Am J Dis Child 1982; 136: 937-41.
11. Friedman SB, Phillips S. What's the difference? Pediatric residents and their inaccurate concepts regarding statistics. Pediatrics 1981; 68: 644-6.
12. Raju TNK, Langenberg PW, Vidyasagar D, Sen AK. A biostatistical survey Questionnaire. J Pediatr 1988; 112: 859-63.
13. Wulff HR, Andersen B, Brandenhoff P, Guttler F. What do doctors know about statistics? Stat Med 1987; 6: 3-10.
14. Cassells W, Schoenberger A, Graboys TB. Interpretation by physicians of clinical laboratory results. N Engl J Med 1978; 299: 999-1001.
15. Norusis MJ. SPSS Inc. SPSS/PC + Base System User's Guide Version 5.0 Illinois : SPSS Inc, 1992: 169-71.
16. Gardner MJ, Altman DG. Statistics with confidence - confidence intervals and statistical guidelines. The British Medical Journal, London, 1989: 72-3.
17. Walter SD. Method of reporting statistical results from medical research studies. Am J Epidemiol 1995; 141: 896-906.
18. Weiss ST, Samet JM. An assessment of physician knowledge of epidemiology and biostatistics, J Med Educ 1980; 55: 692-7.

ความรู้สถิติของแพทย์: การสำรวจในโรงพยาบาลศรีนครินทร์, มหาวิทยาลัยขอนแก่น

*มาลินี เหล่าไพบูลย์, วท.ม., ภิเศก ลุมพิกานนท์, พ.บ.**,*
*Stephen D Walter, Ph.D.****

ปัจจุบันวิวัฒนาการทางด้านการแพทย์ได้ก้าวหน้าอยู่ตลอดเวลา แพทย์จึงมีความจำเป็นต้องติดตามวิวัฒนาการดังกล่าว โดยการอ่านวารสารงานวิจัยทางการแพทย์ และศึกษาค้นคว้าวิจัยในปัญหาต่าง ๆ ที่สนใจ เพื่อให้ทันต่อเหตุการณ์และสามารถนำมาประยุกต์ใช้ในการให้บริการรักษาผู้ป่วยได้อย่างดีที่สุด การจะบรรลุผลดังกล่าวได้อย่างถูกต้อง แพทย์จำเป็นต้องมีความรู้ความเข้าใจในวิธีการสถิติ การศึกษาครั้งนี้ได้กระทำขึ้นเพื่อประเมินความรู้ทางด้านสถิติของแพทย์และนักศึกษาแพทย์ไทย ระหว่างเดือนมีนาคม ถึงมิถุนายน 2535 จากแพทย์ที่ปฏิบัติงานในโรงพยาบาลศรีนครินทร์ คณะแพทยศาสตร์ มหาวิทยาลัยขอนแก่น จำนวน 365 คน ซึ่งประกอบด้วย แพทย์เฉพาะทาง 156 คน แพทย์ประจำบ้าน 152 คน และนักศึกษาแพทย์ปีสุดท้าย (extern) 57 คน โดยใช้แบบสอบถามแบบเลือกตอบซึ่งมีคำอธิบายชี้แจงวัตถุประสงค์และวิธีการตอบให้แพทย์ได้ตอบเอง แบบสอบถามประกอบด้วยข้อคำถามซึ่งมีเนื้อหาเกี่ยวกับการแปลความหมายคำศัพท์ทางสถิติที่พบบ่อยในวารสารงานวิจัยทางการแพทย์จำนวน 9 ข้อ ผู้วิจัยใช้ค่ามัธยฐาน และ 95% CI ของค่ามัธยฐานในการอธิบายคะแนนความรู้ทางด้านสถิติของแพทย์ และเปรียบเทียบความแตกต่างคะแนนความรู้ของแพทย์แยกตามระดับของวิชาชีพ และประสบการณ์เกี่ยวกับการอบรมทางด้านสถิติ โดยใช้การทดสอบแบบมัธยฐานที่ระดับนัยสำคัญ 0.05

ผลการศึกษาพบว่ามีแพทย์ตอบแบบสอบถามกลับมาทั้งสิ้นร้อยละ 40 (146/365) ค่ามัธยฐาน คะแนนความรู้สถิติ เท่ากับ 4.0 คะแนน (95% CI 3.0, 4.0) กลุ่มแพทย์เฉพาะทางมีค่ามัธยฐานคะแนนความรู้สถิติ 4.0 คะแนน สูงกว่าของกลุ่มแพทย์ประจำบ้าน และกลุ่มนักศึกษาแพทย์ ซึ่งมีค่ามัธยฐานเท่ากัน 3.0 คะแนน อย่างมีนัยสำคัญทางสถิติ ($P = 0.02$) กลุ่มแพทย์ที่เคยเข้ารับการอบรมทางด้านสถิติมีค่ามัธยฐานเท่ากัน คะแนนความรู้สถิติ 5.0 คะแนน สูงกว่าของกลุ่มที่ไม่เคยเข้ารับการอบรมทางด้านสถิติซึ่งมีค่ามัธยฐาน 3.0 คะแนน อย่างมีนัยสำคัญทางสถิติ ($P < 0.01$)

จากผลของการศึกษานี้แสดงให้ทราบว่าแพทย์ในโรงพยาบาลศรีนครินทร์มีความรู้ทางด้านสถิติพื้นฐานที่ใช้บ่อยในงานวิจัยทางด้านการแพทย์ไม่เพียงพอ ดังนั้น จึงควรพิจารณาจัดหลักสูตรการอบรมทางด้านสถิติ ให้แก่แพทย์ประจำบ้านและแพทย์เฉพาะทางอย่างต่อเนื่อง

* คณะสาธารณสุขศาสตร์,
** คณะแพทยศาสตร์ มหาวิทยาลัยขอนแก่น, ขอนแก่น 40002
*** McMaster University, Canada.