Biostatistics, epidemiology and vital statistics: an investigation of the difficulty of courses undertaken by medical students in Raipur, India

Swapnil Shinkar [1], Mohd Junaid [2], Sachin Patil [2], Amol R Patil [2], Kalyani Sope [3]*

ABSTRACT

Background

Biostatistics, epidemiology, and vital statistics are an essential part of evidence-based medicine. They are used to generate results and analyze collected data. They are considered to be an important component of research study design for medical professionals. Medical students need to be able to understand such evidence and how it is presented in the literature, as well as be able to make decisions about patient care based on this. Past studies have noted that students find biostatistics difficult when compared with other components of the community medicine module, including epidemiology and vital statistics. The aim of this study was to assess how difficult students find biostatistics, epidemiology and vital statistics, assessing each topic individually and overall, as all three are taught to medical students under the community medicine module in medical schools. Secondly, we assessed if there was any difference in difficulty based on gender.

Methods

A cross-sectional observational study was conducted among 174 undergraduate medical students undertaking external examination, using a set of three questions related to the biostatistics, epidemiology and vital statistics modules. The questions were attempted after 25 hours of teaching. Data were collected and entered in Microsoft Excel and analyzed using IBM SPSS version 23.0 software.

Results

51% of the students were female, 49% were male. Female students correctly answered 46.1% of biostatistics questions, 78.7% of epidemiology questions, and 78.7% of vital statistics questions. Male students correctly answered 40.0% of biostatistics questions, 62.4% of epidemiology questions and 52.9% of the questions on vital statistics. Overall, the male students showed greater difficulty in completing this key community medicine module than females. Both genders found biostatistics more difficult than either epidemiology and vital statistics, but the difference in difficulty was greater for male students.

Conclusion

This study identifies challenges with the teaching of statistical subjects in medical schools which needs to be addressed, particularly with regard to the performance of male students and to the teaching of biostatistics.

Keywords: Biostatistics, Epidemiology, Vital statistics, Gender, Medical schools

INTRODUCTION

Statistics is a science that deals with the collection, classification, tabulation, analysis, interpretation and presentation of numerical data, and the application of this data. When used in the fields of biology and medicine it is known as biostatistics.[1] The role of biostatistics, epidemiology and vital statistics (measurements of the body) is recognized as a core component in the curriculum of undergraduate medical schools and colleges in developed and developing countries.[2] A recent report undertaken in a Malaysian medical institute reported that the value of biostatistics in medicine and healthcare is widely recognized not only by qualified medical practitioners but also the healthcare sector in general.[3] Biostatistics, epidemiology and vital statistics are unpopular subjects in the undergraduate medical curriculum, however, most likely because they encompass mathematics and calculation, which many students find confusing.[3-5]

A previous study published by Verma et al. in 2020[6] rated the difficulty level of biostatistics relative to epidemiology to be 4.8 times higher than the difficulty of epidemiology relative to demography in the pre-university examination. In the final university examination, the difficulty level of biostatistics relative to epidemiology was 7.3 times higher than that of epidemiology relative to demography, suggesting that the problem increases, rather than is addressed by, medical school teaching. This also suggests that even within the broad topic of statistics, some fields of statistics are more difficult, and more difficult for some demographics, than others.

At Ayush Raipur University in Chhattisgarh, India, biostatistics, epidemiology and vital statistics are taught together in third-year MBBS part 1 as a course offering under the discipline of community medicine. The objective of this course is: to enable the medical student to understand the language and principles of biostatistics, epidemiology and vital statistics; to highlight the nature and the distribution of disease; and to teach students to design their research projects, as well as be able to critically read and understand scientific papers in medical journals.

The course covers basic theoretical concepts along with practical sessions. The sequence of the topics is carefully planned so that the three separate components are taught in parallel, which helps students to amalgamate knowledge from each of the topics. Lectures are followed by practical sessions when applicable. In this study, we assess the level of difficulty this key community medicine module poses to medical students in this institute, including to investigate whether a gender difference is observed, as previous studies, e.g. Javali, 2016[7], have found that women score significantly higher in statistical modules than male peers.

The main objective of this study was to assess medical students' performance in epidemiology, biostatistics and vital statistics, to determine the level of difficulty each posed to the students, and to assess the three different topics against one another to assess how difficult the students found one topic compared with the others. The study also assessed whether there was any difference in the difficulty index based on gender.

The module was tested only within a single institute so a limitation of the study is that there is no comparison data available from other institutes; further research is needed to establish whether these results are specific to this institute or generalizable to medical schools across India, South East Asia more widely or internationally.

METHODS AND MATERIALS

A cross-sectional study to record the Difficulty Index of the three topics was conducted at Pt. J.N.M. Medical College, Raipur (Chhattisgarh), after obtaining informed consent from the participants. The study period was from 1 January 2020 to 30 March 2020.

Once a term, medical students are assessed to give insight into their learning and competencies. During such an assessment period, students were assigned three statistical problems to complete under test conditions as part of their routine term assessment. Out of these, one problem was related to biostatistics,

one to epidemiology and one to vital statistics. The problems were administered to 174 study participants who met inclusion criteria for the study (third-year MBBS final part I students who were present in both the pre-university and final university examination of 2020). Exclusion criteria were those who were not present in the pre- or final examination of 2020 or who did not give verbal consent.

We used item analysis to examine student responses to individual test items, to assess the quality of those items and of the tests as a whole, and to indicate where improvement or revisions might be needed.[8-10]

For each problem the proportion of those who could solve it was recorded as a percentage.

P = (No. of individuals answering the problem correctly) / (Total No. of individuals)

Thus, the problems were able to be arranged in order of 'percentage difficulty' – the lower the percentage of the respondent group who correctly answered a particular problem, the more difficult it was considered to be compared with a respondent group

in which a higher score was recorded. Thus, for example, if 8% of male students answered an item correctly but 25% of the female students answered the same question correctly, it would be deemed to be more difficult for males than females.

Statistical analysis
Data entry and analysis was undertaken using Microsoft Excel and IBM SPSS 23.0 to compare the biostatistics, epidemiology and vital statistics modules. Frequencies or percentages are presented for categorical variables.

RESULTS
A total of 174 students participated: 85 (49%) males and 89 (51%) females. All the students were in the 12th class, and 59.2% studied in English (the rest studied in Hindi). The majority of the students were Hindu (67.8%). Only 10.9% students had studied biology with mathematics during their time in 12th class. 78.2% of students attended more than 60 percent of the scheduled classes. The mean age of the participant was 22 years (standard deviation 1.4 years), with a minimum age of 20 years and a maximum age of 30 years recorded (Table 1).

Table 1 Sociodemographic data of study participants

Variables		Number	Percentage
Gender	Male	85	48.9%
	Female	89	51.1%
Religion	Christian	8	4.6%
	Hindu	118	67.8%
	Muslim	1	0.6%
	Other	47	27.0%
12th Class Language	English	103	59.2%
	Hindi	71	40.8%
12th Class Subject	Biology	112	64.4%
	Biology + Mathematics	62	35.6%
Attendance	Less Than 60%	38	21.8%
	More Than 60%	136	78.2%
NEET Exam Attempt	Less Than 2	48	27.5%
	More Than 2	126	72.5%

Out of the total 89 female students, 46.1% correctly solved the biostatistics problem, 78.7% correctly solved the epidemiology problem and 78.7% correctly solved the vital statistics problem in the examination. Out of the total 85 male students, 40.0% correctly solved the biostatistics problem, 62.4% correctly solved the epidemiology problem and 52.9% correctly solved the vital statistics in the examination (See table 2, below).

The difference between the number of female students compared with male students who answered the epidemiology and vital statistics modules correctly was statistically significant (p=0.018 and p=0.00 respectively). While the difference in scores between the female and male students in biostatistics was not significant (p=0.418). Overall the difference in scores for all three modules considered together was statistically significant (p=0.002). (Table 3).

Table 2 Relative difficulty of biostatistics, epidemiology and vital statistics modules

	Biostatistics	Epidemiology	Vital Statistics
Female	46.1%	78.7%	78.7%
Male	40.0%	62.4%	52.9%
% difference M/F	6.1%	16.3%	25.8%

Table 3 Gender distribution of the participants who correctly solved the problems in the biostatistics, epidemiology and vital statistics modules

Module	Gender	Number	Percentage	P value
Biostatistics	Female	41	46.1%	0.418
	Male	34	40.0%	
Epidemiology	Female	70	78.7%	0.018*
	Male	53	62.4%	
Vital Statistics	Female	70	78.7%	0.000 *
	Male	45	52.9%	
Total	Female	23	25.8%	0.002*
	Male	7	8.2%	

*p=0.05 or less deemed to be statistically significant.

DISCUSSION

Quality medical care depends upon the development of knowledgeable, skilled and competent medical personnel. Biostatistics is often considered a difficult subject for medical students, and this study upholds that medical students find biostatistics particularly challenging, with students of both genders scoring on average below 50%. A similar number of male and female participants were selected in this study and women displayed less difficulty with these subjects overall than men based on scores across the three topics (p=0.002). In the cases of epidemiology and vital statistics, the differences in performance between males and females were statistically significant (epidemiology, p=0.018, and vital statistics

p=0.002), though this was not the case in biostatistics (p=0.418), perhaps because both genders scored poorly. This confirms findings from previous studies, e.g. Javali, 2016[7], which also found that women score significantly higher in some statistical modules than male participants. The reasons for this gender difference were not within the scope of this study but invite further research to explore.

No-one can deny the need for these subjects in the medical curriculum. It is generally considered to be a compromise if those in medical research and practice have only basic capability in statistical science. In this regard, the results – especially with regard to

biostatistics, are concerning. So too are the relatively poor scores in epidemiology and vital statistics achieved by male students. This suggests that the way statistics are taught in medical schools needs attention, with better processes introduced to ensure that students are able to follow and retain the teaching they receive.

We suggest that the short duration of subjects at most institutions may prohibit a comprehensive explanation of the topic and may limit students' understanding. The lack of practical exercise as part of the curriculum, for example, presents a major challenge to the students.

CONCLUSION

This study shows that the biostatistics module is more difficult than other modules for all undergraduates and that male undergraduate medical students scored sub-optimally in all statistical subjects. A fundamental understanding of biostatistics, epidemiology and vital statistics is essential for every medical student for understanding the concepts, applications, and importance of many aspects of medicine, public health and demography. If current education curricula are not able to embed this understanding to a sufficient level, new approaches are needed.

RECOMMENDATIONS

1. Every medical institute must incorporate biostatistics and research methodology into their curriculum. This is essential to provide conceptual knowledge. It will also be important to introduce some computerised statistical packages such as Statistical Packages for Social Sciences (SPSS), R Programming, STATA, Python and Microsoft Excel, to enable students to become comfortable with the mathematical calculations and formulae that enable them to them to learn and make use of biostatistics.

2. Increase the number of lectures and appoint additional faculty staff, particularly for biostatistics and research methodology.

3. Scale up this study to investigate the current levels of knowledge and practice in other medical schools, to understand if the difference in the levels of difficulty persist at postgraduate level, and to identify the reasons for the significant gender differences observed in the scores recorded.

ACKNOWLEDGEMENT

REFERENCES

1. Gupta SC, Kapoor VK. Introduction (Meaning & Scope): *Fundamentals of Mathematical Statistics*. 11th rev ed. 2002.
2. Grady CS, Looney Swa, Steiner RP. A Study of Biostatistics in Medical Schools in the United states. In Proceedings of the ASA Sections on Statistical Education: 1994; Alexandria, Virginia, USA ASA; 1994: 263-265.
3. Mostert P. Changing approaches and perceptions: Biostatistics and its role in teaching the Stellenbosch doctor. Proceedings of the seventh International Conference on the teaching of Statistics (ICOTS 7): 2006; Salvador, Brazil Auckland: IASE; 2006.
4. Sahai H. Teaching Biostatistics to medical students and professionals: Problems and solutions. International Journal of Mathematical Education in Sciences and Technology. 1999; 30(2): 187 - 196.
5. Lam YM. Teaching Statistics to Medical Students in the New Medical School. Journal of the Hong Kong Medical Association. 1986; 38(4): 196 - 198.
6. Verma N, Shinkar S, & Jain K. Item test analysis to assess the difficulty index for biostatistics, epidemiology & demography module of undergraduate medical students of Raipur, Chhattisgarh, India. Public Health Rev Int J Public Health Res. 2020; 7(6): 66 - 70.
7. Javali SB, Sunkad M. Comparison of Knowledge and Attitude towards Biostatistics among Oral Health Professionals. Int J Health Sci Res. 2016; 6(9): 32 - 39.
8. Item Analysis. Available from: www.washington.edu./oea/score1/htm.
9. Singh T, Gupta P, Singh D. Test & Item Analysis: Principles of Medical Education. 3rd ed. New Delhi: Jaypee Brothers Medical Publishers ® ltd. 2009; 70-77.
10. Matlock Hetzel S. Basic concept in item & test analysis. Proceeding at the annual meeting of the southwest educational research association. Austin, January 1997.