



LEARNING/TEACHING

Integrating a Primer Course in Biostatistics into the Haematology Practicals of First-Year Medical Students in India

MARIO VAZ

Department of Physiology, St John's Medical College, India

ABSTRACT **Context:** *In India, biostatistics in the medical curriculum is usually taught as a separate course, using either "imaginary" data or data from the literature. Knowledge of statistics is particularly important in the context of "evidence-based medicine".*

Objective: *To evaluate the efficacy of a biostatistics course integrated into the practical haematology first-year medical course with the following special characteristics: (1) students learn statistics on data generated by and on themselves, (2) the course avoids mathematical computation, (3) the statistical exercises are linked to the learning objectives of the physiology curriculum, and (4) the course is without the threat of university examinations.*

Methods: *Statistical exercises were incorporated into specific haematology practicals with the aim of covering simple descriptive and inferential statistics. Statistics tests were administered, without prior information, to 60 first-year medical students before the biostatistics course, immediately following the course, and nine months later. Fifty-four students completed all three evaluations.*

Results: *Students had a poor knowledge of statistics at baseline. They substantially increased their scores in the statistics test immediately following the biostatistics course. These scores remained higher than baseline nine months after completion of the course, although there was a small decline in the absolute scores when compared with scores soon after the biostatistics course.*

Implications: *Integrating biostatistics into other subjects in the medical curriculum may be an important addition to "stand alone" courses in biostatistics.*

KEYWORDS *Humanism, undergraduate medical education, professional development.*

Address for correspondence: Mario Vaz, Department of Physiology, St John's Medical College, Bangalore 560034, India. Fax: +91-80-5532037. E-mail: mariovaz@divnut.net

Introduction

The importance of biostatistics has gained considerable ground with the realisation that medical graduates need to practice evidence-based medicine (EBM) (Bloom, 1995). The underlying skill for EBM is the rigorous application of critical thought and scientific methods to the basic sciences, the social sciences and clinical practice. Acquiring these skills allows students to keep abreast of current literature (Gehlbach, 1988; Linzer *et al.*, 1988) and undertake appropriate health initiatives. The need for a sound knowledge of biostatistics is emphasised by a recent national review on biostatistics requirements in US medical schools (Looney *et al.*, 1998). Sahai and Ojeda (1999) recently outlined the problems of teaching biostatistics to medical students and professionals. Among the problems they cite is the limited time allotted to biostatistics instruction. The problem of limited time for biostatistics could be overcome by integrating biostatistics into other pre-existing courses in the medical curriculum as has been done in the present exercise.

One of the barriers to the study of biostatistics for medical students is their common dislike for mathematics (Vaz *et al.*, 1998) and their perception of statistics as a “mathematical” subject. A second barrier is that the current teaching of biostatistics often draws examples from textbooks rather than the personal experience of students. We have earlier shown that students enjoy the learning process more when they deal with problems related to themselves (Padmavathi *et al.*, 1998).

Addressing these two barriers is important when formulating an appropriate programme in biostatistics. Courses in statistics need to be more conceptual and less mathematical. This approach, in fact, has been recommended for beginners (WHO, 1992). It is also preferable for students to generate data on themselves and then use these data to understand statistics.

The haematology practical course provides an ideal setting for implementing an integrated biostatistics programme, where the data that students generate on themselves can be subjected to statistical analysis for further interpretation. The programme outlined here was developed in our institution. We present an evaluation of the short- and long-term impact of this programme on our students’ statistical knowledge.

St John’s Medical College is situated in Bangalore, South India. This school’s curriculum is essentially that dictated by the Medical Council of India. Under this system, human physiology is a 1-year course and practical haematology is assigned a total of 30 hours.

Methods

Sixty students admitted into the 1998 MBBS (Bachelor of Medicine and Surgery) batch of the medical school underwent the biostatistics programme

during their first year in medical school. Students admitted into the medical school come from various states in India, only some of which require statistics to be taught at the higher-secondary level prior to admission into medical school. To evaluate statistical knowledge at baseline, all the students were administered a statistics pre-test on admission to the medical school. This test, which did not require any mathematical computation, mainly required students to “fill in the blanks”. The questions covered a range of topics including measures of central tendency and dispersion, the normal distribution, the interpretation of p value and the interpretation of correlation coefficients. The test had a maximum of 15 marks of which seven were allotted to descriptive statistics, seven to inferential statistics and one to experimental design. This test was administered again at the end of the practical haematology course and at the end of the first year of medical school, approximately 9 months after they completed the practical haematology course. The investigator formulated the statistics test to assess a broad knowledge of basic statistics, which was the focus of the integrated biostatistics course.

A control group, either not receiving any training in biostatistics or receiving a traditional biostatistics course, would have allowed for a more stringent evaluation of the present integrated biostatistics course. Neither was possible, however, because of practical constraints in accommodating such a study design within the existing schedule of courses. A quasi-experimental design was adopted, in which statistical knowledge was evaluated over time with each student acting as his/her own control.

Biostatistics is not taught as part of the curriculum to first-year medical students. The framework of the present biostatistics course was developed on the basis of experience gained by the author while conducting yearly introductory workshops on biostatistics for medical graduates. Details of this biostatistics course, which was integrated into the haematology practical course of first-year medical students, are given in Table 1. For each of the practicals indicated in Table 1, students generated data on themselves. Thus, statistical analyses generally involved a sample size of 60. At the end of each practical, students noted their data on the blackboard. Students were encouraged to speculate on what the collective data might represent prior to any statistical analysis. These data were then analysed by the investigator and the descriptive and inferential statistics derived from these data were placed on the notice board a day later. The students then transcribed these statistics onto their practical worksheets and interpreted the analysed data. The tutors corrected the practical worksheets at the next practical class. The emphasis during the course was, thus, not on mathematical computation (which was done by the investigator), but on interpreting data, such as the student might encounter while reading medical literature. As part of the biostatistics course, after the first haematology practical, the students had a single 60-minute introductory lecture on statistics and were given a three-page handout that they were encouraged to refer to for the rest of the course. This introductory lecture

Table 1. An outline of the biostatistics programme that was integrated into the Haematology Practical Course

Practical	Statistical exercise	Learning outcomes
Estimation of haemoglobin (Hb) (whole batch works on a standard sample)	<ul style="list-style-type: none"> ● Mean, SD, coefficient of variation (CV) 	<ul style="list-style-type: none"> ● Extent of errors associated with laboratory tests
Estimation of haemoglobin (each student estimates their own sample)	<ul style="list-style-type: none"> ● Comparison of data from males and females in the class by “GROSS” means, i.e. mean, SD, frequency distribution ● Mean, SD, CV 	<ul style="list-style-type: none"> ● Reporting descriptive data ● What is a “normal” distribution?
RBC count (whole batch works on a standard sample)	<ul style="list-style-type: none"> ● Comparison of CVs obtained for Hb estimation and RBC count 	<ul style="list-style-type: none"> ● Further understanding of errors in laboratory tests
RBC count (each student estimates their own sample)	<ul style="list-style-type: none"> ● Compute the male and female data separately (mean, SD) ● Comparison of gender differences using inferential statistics (independent <i>t</i>-test) ● Reanalyse the Hb gender data using independent <i>t</i>-test 	<ul style="list-style-type: none"> ● Why do different analytical tests have different CVs? ● Importance of inferential statistics
Practical statistical exercise	<ul style="list-style-type: none"> ● Generate a “scatter plot” for Hb vs RBC count for the whole class 	<ul style="list-style-type: none"> ● What is the Null hypothesis? ● How do you interpret a <i>P</i> value?
Total WBC count (each student does his own sample)	<ul style="list-style-type: none"> ● Calculate the “range” of the data for the entire class ● What is the 50th percentile for the entire class? 	<ul style="list-style-type: none"> ● Can “observed” differences be misleading? “Observed” and “statistical” difference ● What is a “positive” and “negative” correlation? ● Other ways to describe data
Platelet count, bleeding time (each student performs the tests before and 2 hours after the drug; half the class gets aspirin 500 mg, half the class gets a placebo)	<ul style="list-style-type: none"> ● Describe the data of the whole class in relation to bleeding time before and after the drug 	<ul style="list-style-type: none"> ● Range, centiles, tertiles, quartiles, quintiles, 50th percentile ● Introduction to paired analyses: the paired <i>t</i>-test ● Comparison of independent and paired analysis ● What is “bias”, “placebo” effect, and “control” group?

included very basic information on measures of central tendency, dispersion, the meaning of a p value, the meaning of positive and negative correlation, and the meaning of the null hypothesis.

Prior to the course, students were informed that the statistical exercises were not a part of the university curriculum and that they would not be tested on statistics at the university examinations in physiology. Statistical exercises were not included in all the practicals since some of the faculty of the Department of Physiology were concerned that the biostatistics programme would overshadow the primary objectives of the haematology course. For this reason, both post-tests in statistics were administered without prior announcement and the performance of students in these tests, therefore, represents knowledge acquired without specific preparation for an examination.

The marks achieved by the students in the statistics pre- and post-tests are presented as mean \pm standard deviation (SD). A comparison of marks across the three tests was done using a Repeated Measures Analysis of Variance (ANOVA) with *post-hoc* testing (paired t -test with Bonferonni correction). The null hypothesis was rejected at $p < 0.05$.

Results

There were important statistical lessons that emerged during the integrated biostatistics course. During the first practical, all students estimated haemoglobin concentrations on a single sample of blood using Sahli's method. This allowed them to assess the between-student variation in assays, and allowed for emphasis of technical errors during the estimation, as well as the need for quality control. The coefficient of variation for the RBC count (33.8%), again determined from estimations done of a single sample of blood, was considerably greater than that for haemoglobin estimation (16.0%). This allowed for a discussion of the complexity of assay procedures on assay variations. The comparison of haemoglobin and the RBC counts between the genders allowed for a much fuller discussion of gender differences since, despite the technical errors made by the students, the differences between the two groups were statistically significant (haemoglobin g%; males 15.0 ± 1.8 vs females 12.0 ± 2.4 , $p < 0.05$; RBC count million/cu.mm; males 5.0 ± 1.2 vs females 4.2 ± 1.2 , $p < 0.05$). A correlation between RBC count and haemoglobin helped the students understand the meaning of correlations. The study of bleeding time before and two hours after the administration of aspirin allowed for the statistical analysis of paired data. The use of a placebo allowed for the emphasis of "bias" and "placebo" effect. In the aspirin group, bleeding time increased significantly by 34.6 ± 27.7 secs ($p < 0.05$). In the placebo group, bleeding time also rose significantly by 14.8 ± 21.0 secs ($p < 0.05$). The increments in the bleeding times of the two groups were significantly different from each other ($p < 0.05$).

The data of the three statistics tests, prior to, immediately following, and nine months after the integrated biostatistics course are presented in Table 2. Before the course, students scored poorly in the pre-test, with their entire knowledge of statistics being limited to descriptive statistics. Immediately following the course, there was a significant rise in the marks obtained in the statistics test. In absolute terms, the students scored higher marks in the descriptive statistics section than in the inferential statistics section. Nine months after the course, with no specific attempts at reinforcement, there was a small decline in the total marks achieved in the statistics test. However, the marks obtained both in descriptive and inferential statistics were still higher than the initial scores obtained in the pre-test.

Fifty-nine out of 60 students were present during a formal assessment of the integrated biostatistics course. Three of these students did not provide any feedback. Of the remainder, 52 (out of 56) felt that the emphasis of biostatistics during the haematology course was “just right”, three felt it was “too little” and one felt it was “too much”. All 56 students who provided feedback felt the course should be continued for future students. A total of 24 students provided suggestions for the improvement of the biostatistics course. Among the suggestions were the provision of revision statistical exercises ($n=5$) and statistical exercises from literature ($n=3$) to provide a clearer understanding of the “real-life” use of statistics for doctors. Two students felt there should be greater emphasis on mathematical computation and two suggested group project work that would call for the use of statistics.

Table 2. A comparison of statistical test scores before and following the integrated biostatistics programme

	Prior to the biostatistics course	Immediately following the course	Nine months after the course
Descriptive statistics (max=7)	1.39 ± 0.86 (1, 0–3)	5.17 ± 1.71* (5, 1–8)	4.95 ± 1.58 (5, 1–7)
Inferential statistics (max=7)	0 (0, 0–0)	3.03 ± 1.84* (3, 0–7)	2.29 ± 1.65+ (2, 0–7)
Total score ^a (max=15)	1.54 ± 1.02 (1.5, 0–4)	8.63 ± 3.45* (8, 2–15)	7.78 ± 3.22 (8, 2–14)

Data are mean ± SD. Figures in parentheses indicate median and range. The data presented are of 54 students, who took part in all three evaluations. Repeated measures ANOVA indicated that there was a significant change in the parameters across the three occasions ($p < 0.01$). A paired t -test was applied for comparison between “prior to” and “immediately following” ($*p < 0.05$), and “immediately following” to “9 months after” ($+, p < 0.01$). The statistical significance indicated in the table is after the Bonferonni correction.

^aThe total score is the sum of descriptive statistics + inferential statistics + 1 mark on experimental design.

Discussion

The present exercise indicates that biostatistics can be successfully integrated into a haematology practical course. In addition, the statistics topics that can be covered are fairly diverse. In a broader context, the implementation of this course is important because there is a need to integrate across individual subjects, with the aim of enhancing the overall skills of medical graduates. In this context, physiology can play an important role in enhancing critical scientific reasoning (Monekosso, 1998), of which biostatistics is an essential component.

The actual marks obtained in the statistics tests following the course would have been higher if prior notice were given to students. However, the strength of the present exercise is that despite the absence of student preparation, there was a substantial increase in statistical knowledge. A further strength is that this increase occurred despite informing students that the statistics course was not a formal part of the curriculum or the examination structure. Earlier studies in our department had suggested that the concerns of students are particularly linked to examinations (Vaz, 1999).

There was a preferential gain in descriptive statistical knowledge as a result of the course. We believe that the relatively modest gain in knowledge of inferential statistics was related, in part, to the fact that this was a new area for all students. In addition, the course did not allow for repetition of concepts. There is, thus, a need to integrate biostatistics courses, similar to that which we employed, into other courses at various stages of the medical curriculum. This will allow for the acquisition and reinforcement of a substantial amount of knowledge of statistics. The integration of biostatistics need not interfere with the primary course. In the current example, biostatistics enhanced the primary course by highlighting phenomena that would have otherwise only been “theoretical”. In addition, the feedback from students indicated that the emphasis on biostatistics during the haematology course was “just right”.

In our exercise the course was formulated and given not by a biostatistician but by an individual more aware of the practical uses of statistics than of its mathematical basis. The advantage is that such a course is likely to be less “technical”. The problem, however, is that the approach and emphasis on biostatistics is likely to be considerably more varied, depending on the competence and motivation of the person giving the course.

Integrated biostatistics courses, such as described here, can help to introduce students to the importance of biostatistics in medicine, and add to the learning outcomes of specialised biostatistics courses within the curriculum. Sahai and Ojeda (1999) stress that a realistic approach to biostatistics should, among other things, avoid excessive mathematical computations, be realistic and allow for the interpretation of “real-life” situations. Our integrated biostatistics course has, we believe, provided a realistic approach to biostatistics teaching and could provide a framework for similar innovations in other schools and other areas of the medical curriculum.

Conclusion

In conclusion, we have described a successful implementation of a biostatistics course, integrated into the haematology practical course of first-year medical students. The course resulted in a significant gain in statistical knowledge and was not perceived as an additional burden by the students who, to a large extent, considered the course content as “just right”. This paper provides the framework for course coordinators in medical and other health professions curricula to think of other situations in which biostatistics can be integrated. This is clearly important, since the current teaching of biostatistics appears to be far less than what is required for medical graduates to effectively practice evidence-based medicine.

Acknowledgements

This study was funded by the Research Society, St John’s Medical College. The author would like to thank the faculty of the Physiology Department who were involved in the conduct of the haematology course for their cooperation.

References

- BLOOM, S.W. (1995). The place of science in health professions. *Medical Education*, 29, 76–78.
- GEHLBACH, S.H. (1988). *Interpreting the medical literature. Practical epidemiology for clinicians*, 2nd edn. New York: Macmillan.
- LINZER, M., BROWN, J.T., FRAZIER, L.M., DELONG, E.R. & SIEGEL, W.C. (1988). Impact of a medical journal club on house-staff reading habits: knowledge, and critical appraisal skills. *Journal of the American Medical Association*, 260, 2537–2541.
- LOONEY, S.W., GRADY, C.S. & STEINER, R.P. (1998). An update on biostatistics requirements in U.S. medical schools. *Academic Medicine*, 73, 92–94.
- MONEKOSSO, G.L. (1998). Essential medical studies: introducing a core curriculum based on the doctor–patient encounter as a catalyst for medical education reform. *Medical Teacher*, 20, 536–543.
- PADMAVATHI, R., MARUTHY, K.N., BOROGHONA, S. & VAZ, M. (1988). The perceptions of first-year medical students on animal and human experiments in physiology. *Indian Journal of Physiology and Pharmacology*, 42, 127–130.
- SAHAI, H. & OJEDA, M.M. (1999). Problems and challenges of teaching biostatistics to medical students and professionals. *Medical Teacher*, 21, 286–288.
- VAZ, M. (1999). What do students expect of lecture handouts during the first year of the medical curriculum? *Medical Teacher*, 21, 324–325.
- VAZ, M., PADMAVATHI, R. & VEENA, G. (1998). Integrating biostatistics and experimental physiology: student’s perceptions and future scope. *Indian Journal of Physiology and Pharmacology*, 42, 453–459.
- WORLD HEALTH ORGANIZATION (WHO) (1992). *Health research methodology. A guide for training in research methods*. New Delhi: Oxford University Press.