Data Entry Skills in a Computer-based Spread Sheet Amongst Postgraduate Medical Students: A Simulation Based Descriptive Assessment

Amir Maroof Khan, Dheeraj Shah¹, Pranab Chatterjee
Departments of Community Medicine, and ¹Pediatrics, University College of Medical Sciences, New Delhi, India

ABSTRACT

Background: In India, research work in the form of a thesis is a mandatory requirement for the postgraduate (PG) medical students. Data entry in a computer-based spread sheet is one of the important basic skills for research, which has not yet been studied. This study was conducted to assess the data entry skills of the 2nd year PG medical students of a medical college of North India.

Materials and Methods: A cross-sectional, descriptive study was conducted among 111 second year PG students by using four simulated filled case record forms and a computer-based spread sheet in which data entry was to be carried out. Results: On a scale of 0-10, only 17.1% of the students scored more than seven. The specific sub-skills that were found to be lacking in more than half of the respondents were as follows: Inappropriate coding (93.7%), long variable names (51.4%), coding not being done for all the variables (76.6%), missing values entered in a non-uniform manner (84.7%) and two variables entered in the same column in the case of blood pressure reading (80.2%). Conclusion: PG medical students were not found to be proficient in data entry skill and this can act as a barrier to do research. This being a first of its kind study in India, more research is needed to understand this issue and then include this yet neglected aspect in teaching research methodology to the medical students.

Keywords: Computer, data entry, medical education, medical postgraduate, spread sheet

Introduction

Doing a research work for a thesis is mandatory for postgraduate (PG) medical students in India.[1] Computers and softwares have become indispensable tools for research nowadays. Research has been conducted on the medical students using computers nowadays, which has found that most of them are using it for socializing or for other non-academic purposes.[2] It has also been reported that most of the research done by the PG medical students is quantitative.[3] A seemingly small but a vital component of the process of conducting a research is data entry. Computer-based data entry is the preferred modality nowadays as it paves a smooth transition for data analysis using computer-based statistical softwares. If data entry is done properly, it can facilitate data analysis to a large extent. A properly done data entry is referred to as “tidy data” which is defined as easy to manipulate, model and visualize and has a specific structure: Variables are stored in columns and observations in rows and one type of experimental unit per file.[4] If the textual responses are pre-coded into numerical codes during the data entry itself; carrying out complex statistical analysis like building regression models becomes less time consuming later. Data entry in a format, which cannot be handled by the statistical software being used, leads to loss of man-hours and also puts mental stress on the researcher at that stage. Research on computer skills have been done in India, but research on how data entry is being done by the PG students, after collecting the data from the study subjects, has not been found in published literature.[2,5]

Though the goal of post-graduation courses is to create specialists, the need of the hour for India is to create physician-researchers who can transition between specialties in order to translate information to best serve the needs of their patients. With the expansion of the horizon of the medical sciences, it has become imperative that even specialist physicians should be fluent as generalist researchers. In this environment, this study looks at the budding specialists and their ability to conduct clinical research,
specifically, at their skills in entering data into structured spread sheets. This issue assumes particularly important proportions for the primary care physician, who is uniquely positioned at the junction of all the specialties and deals with the whole spectrum of problems that the patient presents with. Thus, primary care physicians in training need to be not only conversant in the jargon of various specialties, but also needs to have a good grasp over the basics of medical research. This study aims to fulfill the lacunae in the evidence-base regarding the skills of physicians that goes in shaping the scene behind the published research.

Objective
To assess the data entry skill in a computer based spreadsheet among the 2nd year medical PG students.

Materials and Methods
A descriptive, cross-sectional study was carried out among the 2nd year medical PG students, who had either completed or were on the verge of completion of data collection at the time of this study. Data was collected during an annual workshop on “thesis writing” for these students. The PG students were given four simulated filled case record forms and were asked to do the entry of this data in a blank spread sheet as is the format of Microsoft Excel that they most commonly used for their data entry. Each simulated case record form was pre-filled with the information related to individuals who were presumed to have been enquired for certain particulars related to socio-demographic characteristics, clinical measurements such as blood pressure, body mass index and some investigation reports such as fasting blood sugar level, etc., As entering missing values is also a typical sub-skill, some responses were kept as blank, to see how the PG students enter the missing values in the spread sheet. There were in all ten sub-skills that were assessed related to the computer-based data entry skill wherein each sub-skill which if found to be correct was given a score of one and a score of zero was given for an improper data entry regarding each specific sub-skill. Thus, the total score ranged from a minimum of zero to a maximum of ten.

Statistical analysis
Data was entered in MS Excel and R software was used for preparing one-way tables and data are presented in the form of percentages.

Results
In all, data entries by 111 PG students were assessed. It was found that most students had scored between 5 and 7 (61.3%), followed by a score between 0 and 4 (21.6%) and only 17.1% scored above 7. Table 1 shows the proportion of students who did the data entry in a proper manner with regards to the specific sub-skill of data entry. The specific sub-skills that were found to be lacking in more than half of the respondents were as follows; inappropriate coding (93.7%), long variable names (51.4%), coding not being done for all the variables (76.6%), missing values entered in a non-uniform manner (84.7%) and two variables entered in the same column in the case of blood pressure reading (80.2%).

Discussion
This study assessed the data entry skill in a spread sheet among the medical PG students in a medical college. A large gap has been identified in computer-based data entry skill using a simulated assessment in this cross-sectional study. The students were lacking in almost all the sub-skills that are required to do a proper data entry for further data analysis in health research.

Rowe et al., in their study have reported from Canada that only 15.0% of the 1st year family medicine residents had performed spread sheet applications.\(^{[8]}\) Only 40.3% of 1st year medical students in a study from Sri Lanka have reported using computers for data base management.\(^{[7]}\) A study from South India has reported that only 23.2% of the medical students surveyed were “very confident” in using spread sheets.\(^{[6]}\) A study from Pennsylvania has assessed the spread sheet and data management skills of students and found that their elementary Microsoft Excel skills were very limited and they have suggested the need to change their curriculum for teaching them.\(^{[5]}\) It is clear that students belonging to different courses need to possess the relevant course-specific skills regarding the use of a particular computer application.

Table 1: Frequencies of correct responses for the items used for assessing the data entry skill of the postgraduate medical students (n=111)

| Items used for assessing the data entry skill | Correct responses n (%) |
|---------------------------------------------|------------------------|
| Variables in columns and records in rows    | 105 (94.6)             |
| One variable per column                     |                        |
| SBP/DBP entered in two columns              | 22 (19.8)              |
| Age/sex entered in two columns              | 82 (73.9)              |
| Units not mentioned alongside the values in the cells | 79 (71.2) |
| Missing values left blank or uniform coding done for it | 17 (15.3) |
| Coding done for all the variables           | 26 (23.4)              |
| Uniformity of coding followed throughout    | 87 (78.4)              |
| Variable names written in a concise format  | 54 (48.6)              |
| Attempt made to enter data related to variables relevant to the objective of the research example given | 73 (65.8) |
| A long response (address) was coded according to the objective of the research example given | 7 (6.3) |

\(^{[5]}\)SBP/DBP: Systolic blood pressure/Diastolic blood pressure
the data that will be generated at this level will have the potential to guide the primary and secondary level prevention activities which needs to be put in place for reducing the disease burden. Through larger, multi-centric studies, we need to examine whether this trend is true across a larger geographical stretch or is a local anomaly. If it is indeed a trend across multiple, representative centers, then we would need to take a relook at the training curriculum for all PG students, with special emphasis on the curriculum for primary care and family physician training programs to ensure that enduring research skills are inculcated during their PG training program.

Limitations
The present study did not collect data regarding the specialties of the residents, mainly because the PG students had entered the specialty courses after completing their medical undergraduate course and as this training is uniform, irrespective of which specialty a student's opt for his/her PG course. Furthermore, they had hardly spent around 2-3 months in the respective department after which this assessment was carried out just before their first formal training in “Thesis protocol writing.” Hence, there was no reason to assume that specialty must have made any difference to their data entry skills at this early stage.

Conclusions
A sizeable proportion of the PG students studied did not know of the various basic aspects of doing a proper data entry in a spread sheet. Capacity building with the specific and relevant skill based training is urgently required for the PG medical students, as they are the doctor, academicians and researchers of the future.

Recommendations
Inclusion of a session on “data entry in a spread sheet” in the research methodology training workshops is the need of the hour for the PG students especially belonging to family and primary care physicians and other research intensive PG programs like Community Medicine.

References
1. Post graduate medical education regulations 2000, Amended upto May 2013, Medical Council of India. New Delhi: Medical Council of India. Available from: http://www.mciindia.org/RulesandRegulations/PGMedicalEducationRegulations2000.aspx. [Last cited on 2013 Nov 18].
2. Maroof KA, Parashar P, Bansal R. How are our medical students using the computer and internet? A study from a medical college of north India. Niger Med J 2012;53:89-93.
3. Dhaliwal U, Singh N, Bhatia A. Masters theses from a university medical college: Publication in indexed scientific journals. Indian J Ophthalmol 2010;58:101-4.
4. Wickham H. Tidy data, 2013. Available from: http://www.vita.had.co.nz/papers/tidy-data.pdf. [Last cited on 2013 Nov 19].
5. Seetharaman N. Assessing computer skills among graduate medical students in south India. Natl J Res Community Med 2012;1:26-34. Available from: http://www.medjourn.in/pdf/Individual%20Full%20Articles/5Dr.Seetharaman_Full%20Article.pdf. [Last cited on 2013 Nov 19].
6. Rowe BH, Ryan DT, Therrien S, Mulloy JV. First-year family medicine residents’ use of computers: Knowledge, skills and attitudes. CMAJ 1995;153:267-72.
7. Ranasinghe P, Wickramasinghe SA, Pieris WA, Karunathilake I, Constantine GR. Computer literacy among first year medical students in a developing country: A cross sectional study. BMC Res Notes 2012;5:504. Available from: http://www.biomedcentral.com/1756‑0500/5/504/. [Last cited on 2013 Nov 19].
8. Baugh JM. Assessment of spreadsheet and database skills in the undergraduate student. Inf Syst Educ J 2004;2:1-12. Available from: http://www.isedj.org/2/30/ISEDJ.2(30).Baugh.pdf. [Last cited on 2013 Nov 19].

How to cite this article: Khan AM, Shah D, Chatterjee P. Data entry skills in a computer-based spread sheet amongst postgraduate medical students: A simulation based descriptive assessment. J Fam Med Primary Care 2014;3:216-8.

Source of Support: Nil. Conflict of Interest: None declared.