Abstract
In any academic field the question of quality is central issue. Evaluation of the quality in supervised research in medical education is very important in higher education and teaching. The purpose of the supervised research project is to provide the learner an experience in scientific writing, to familiarize the student with the steps of research, to promote critical and analytical thinking of the learner and to contribute something new to knowledge base. The objectives of this study were to determine the quality in postgraduate supervised research report writing, according to the standard recommended and implied by the post graduate medical institutes of Bangladesh and thus to assess the achievement of the intended purposes. For that an exploratory and descriptive study was performed on 64 theses and dissertations submitted during the period of 2008-2010 in two selected post graduate medical institutes of Dhaka city. The overall quality was found marginal in 93.8% (60) of the reports. There were indications that the prescribed standard was not followed properly. It was observed from the study that there are problems in learning and inadequacy in fulfillment of the purpose of supervised research program. It was perceived that probably the supervision was not enough. Theoretical knowledge gap was also apparent at various levels. To address these issues planning, designing, supervision and assessment systems of supervised research program in relation to the postgraduate medical qualifications of Bangladesh needs critical attention.

Key Words: Research, Post graduate medical education, Quality

Introduction
Research is believed to be the way to educate the learners for supporting and extending the knowledge base applicable to the rapidly changing environment. Medical practitioners' needs 'professional wisdom' for best 'professional judgment'. The exercise of academic research is to learn, build up and develop 'professional wisdom' which is a complicated amalgam of factual knowledge, procedural skill and intuition. Writing thesis/dissertation is to be familiar with research process, is to acquire scientific writing skill and also for 'learning to learn'.

In the scientific community there is a belief that the qualities of the medical thesis/dissertation are less than what is expected in the current context. The process of evaluation is an integral part of any educational program. Researches focusing on the quality of thesis/dissertation in higher education are receiving little interest and priority. Thesis or dissertation is document of the candidate's research work, which reflects their learning outcome. Therefore to assess the learning quality; the thesis/dissertation assessment can be carried out.

The aim of this study was to provide complimentary information to support the institutes to take steps for further improvement. It was expected that this work would be able to identify the important areas of weaknesses and areas for further research. It was also predicted that the finding of this study will be able to help the learners to attain their learning objectives, will assist the supervisors to concentrate on the critical areas for supervision and will make the jobs easier for the examiners.

After extensive literature search, institutional records and personal communication with the resource personnel, the investigators have failed to locate any evidence or documents on the study of quality issue of thesis/dissertation in any institute of Bangladesh. Therefore this study is to be considered as the beginning of this type of works.

1Professor, General Surgery, Head of the department, Anwer Khan Modern Medical College, Dhanmondi, Dhaka
2Professor, Curriculum development, and course coordinator, MMEd, CME, Mohakhali, Dhaka
3Assistant Prof. Dermatology and Venereology, OSD DGHS, Mohakhali, Dhaka
4Lecturer, Chemistry of Dental Material, CME, Mohakhali, Dhaka
5Associate Professor, Gynecology and Obstetrics, Mymensingh Medical College
6Professor Orthopedic Surgery, Barisal Medical college, Barisal
7Associate professor, Ophthalmology, OSD DGHS, CME, Mohakhali, Dhaka
8Consultant, Orthopedic Surgery, Singair, Manikganj
9Professor, Department of Pediatric Surgery; Medical Educationist, Bangabandhu Sheikh Mujib Medical University, Dhaka

Address of correspondence: Professor Dr. Md. Abdul Wohab Khan, Department of surgery, AKMCC, Dhanmondi, Dhaka. E mail dhhkhan1962@live.com Cell: +88-01711-52-1064

Bangladesh Journal of Medical Education 2012;3(1):1-7. © 2012 Khan et al., publisher and licensee Association for Medical Education. This is an Open Access article which permits unrestricted non-commercial use, provided the original work is properly cited.
Methodology
This was an exploratory, cross sectional and descriptive study. The study was done at Centre for medical education (CME) Bangladesh. Theses/dissertations submitted and accepted through January 2009 to December 2010 and available in the library were included. The document not available in the library and those used for pretesting were excluded.

Total 64 reports were reviewed with a convenient sampling of the institutes and departments; all available (31) theses were taken from one institute. Rest was taken from the other institute with systematic sampling. Four instruments were developed and used for measurement and collection of data. The 'rating scale' was created to rate on the quality of the variables supported by a 'rubrics'. A 'comprehensive rating scale' was developed to calculate the overall grading. A separate 'checklist' was compiled to enlist the weaknesses found in sections. The sections of a thesis/dissertation (from title to annexure), General organization and physical format, innovation and creativity were examined separately as the variables. The 'statements to define the criteria (rubrics)' were collected and rephrased to increase objectivity. The rubrics were clustered into four groups to describe the grades. A score of 80 and above out of 100 was branded as 'exemplary', 60 to 79 as 'good' and 50 to 59 as 'average'. Less than 50 score was categorized as 'marginal'. Marks were allocated to each of the criteria according to weightage assuming its role to achieve the learning objectives. An allowance of total 100 (one hundred) was considered for individual section for convenience of scoring. The comprehensive rating scale for total document (thesis) was framed on 100 (one hundred) marks. Allocation of marks to each of the sections was specified according to its importance to achieve the educational goal described in the institutional guideline.\textsuperscript{5,6,8} The instruments were developed and pretested several times before finalization. Every time, the pretesting was done on a selected sample of thesis by random sampling from study population and each thesis was examined by one of the senior examiners and also by the investigator himself to cross check the inter rater reliability.

The process of document analysis and review was adopted for data collection using and supported by the prepared rubrics. The whole document was reviewed repeatedly to check and recheck for rating, qualitative grading and to gather informations on mistakes and weaknesses. An overall rating was calculated and recorded in the comprehensive rating scale as per the assigned weightage. Data were labeled, coded and entered manually in SPSS version 15.11 for Windows 7 for analysis.

The study was a document review and do not involve any animate subject. To collect the necessary information and data, due permissions from the institutional or departmental authority were taken; mentioning the purpose of study. No identity of the institute or any person was disclosed to ensure confidentiality.

Results
The distribution of marks in different sections
Minimum numbers obtained in sections were zero and maximum was 75. The highest mean score (49.87) with minimum dispersion (SD 8.78) was in the title section. Maximum dispersions were noticed in the recommendation, Appendices and abstracts (Table 1).

Table 1: Distribution of marks scored by different sections of the reports under study (n = 64)

<table>
<thead>
<tr>
<th>Sections</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Title</td>
<td>25</td>
<td>68</td>
<td>49.87</td>
<td>8.782</td>
</tr>
<tr>
<td>Abstract</td>
<td>0</td>
<td>70</td>
<td>31.63</td>
<td>18.385</td>
</tr>
<tr>
<td>Introduction</td>
<td>18</td>
<td>63</td>
<td>35.33</td>
<td>10.794</td>
</tr>
<tr>
<td>Literature review</td>
<td>0</td>
<td>60</td>
<td>36.98</td>
<td>15.694</td>
</tr>
<tr>
<td>Methodology</td>
<td>18</td>
<td>75</td>
<td>46.70</td>
<td>11.200</td>
</tr>
<tr>
<td>Results</td>
<td>25</td>
<td>65</td>
<td>43.78</td>
<td>8.207</td>
</tr>
<tr>
<td>Discussion</td>
<td>0</td>
<td>65</td>
<td>40.50</td>
<td>10.044</td>
</tr>
<tr>
<td>Conclusion</td>
<td>0</td>
<td>60</td>
<td>38.13</td>
<td>13.437</td>
</tr>
<tr>
<td>Recommendation</td>
<td>0</td>
<td>50</td>
<td>19.69</td>
<td>20.332</td>
</tr>
<tr>
<td>Citation</td>
<td>0</td>
<td>75</td>
<td>49.83</td>
<td>10.672</td>
</tr>
<tr>
<td>initiative/creativity</td>
<td>0</td>
<td>50</td>
<td>6.09</td>
<td>12.550</td>
</tr>
<tr>
<td>Appendices</td>
<td>0</td>
<td>60</td>
<td>38.83</td>
<td>20.173</td>
</tr>
<tr>
<td>General qualities</td>
<td>20</td>
<td>59</td>
<td>43.72</td>
<td>8.620</td>
</tr>
</tbody>
</table>

The distribution of the quality grades appraised for individual sections
In all of the sections, marginal quality was found in majority of the cases (39% to 98%). The sections with lower percentages of marginal grade were the title (25, 39%), literature review (39, 61%) methodology (36, 56%) citation (30, 47%) and appendices (30, 47%). Maximum 'marginal' worth was (63, 98%) found in innovation and creativity. Other sections with high grade of marginality were abstract (56, 87%), introduction (56, 87%), discussion (54, 84%), conclusion (53, 83%) and recommendation (55, 86%). None of the sections in any document out of 64 could be graded as exemplary (Fig 1).
The distribution of marks and quality of the total documents

Mean of the calculated score for the whole document was 40.73 (SD 7.27). The median was found 41 and smallest mode was 38 (Table 2). The calculated grading of the total document was appraised as marginal in 60 (93.8%) and average in 4 (6.3%) of the documents examined (Figure 2).

Table 2: The mean median and mode of overall marks calculated for the whole document (n = 64)

<table>
<thead>
<tr>
<th>Statistics</th>
<th>Comprehensive total marks</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>40.73</td>
</tr>
<tr>
<td>Median</td>
<td>41.00</td>
</tr>
<tr>
<td>Mode</td>
<td>38(a)</td>
</tr>
<tr>
<td>SD</td>
<td>7.269</td>
</tr>
</tbody>
</table>

Overall grading of the reports as appraised by the comprehensive rating calculated from the marks obtained in the rating scale

The calculated grading of the total document was appraised as marginal in 60 (93.8%) and Average in 4 (6.3%). None of the reports was good or exemplary.

Common weakness and pitfalls identified in different sections

The phrasing in many titles was confusing and some was self-contradictory and most of them were not clear enough to communicate the topic. Most common shortcoming in abstract was a missing or unclear research question. Both summary and abstract were added in a small number of cases and in few documents it was missing. Absence of operational definition, missing list of variables or inappropriate variable list was common mistakes found in the introductory chapter. Failure to identify a knowledge gap or to translate a problem was comprehended in many. Presence of irrelevant information was identified in most of the descriptions. Other uncommon flaws were absence of true hypothesis or research question, improperly formulated objectives.

Literature review contains informations which were not relevant with the topic in most of the report examined. Many failed to describe any recent thought. The literature review was absent in a few of the Thesis/ dissertations. Most consistent shortcomings in the methodology sections were inadequate or absence of description on sample size estimation and sampling techniques. Inappropriate choice of research design to solve the selected problem was another weakness found in some cases. Limitations were not defined in many. Other drawbacks were incomplete description of the data collection process.
Data presentation was not focused to answer the research question; was the most frequent weakness. The next common fault was failure to arrange the data pointing to measure the objectives. In some of the reports tables and figures were presented with incomplete format or both were representing the same data. Most of the discussions failed to explain the implication of result to answer the research question rather they were mere repetitions of the result. The implication of limitations on the result or conclusion was not echoed in many cases. Great majority of conclusions were supported by very weak evidences or based on evidences from others. Some conclusions were mere repetition of the former sections. The recommendation was absent in some of the reports. It was hard to consider them as valid in many cases. In majority of the reports institutional format style was not followed correctly in citing references. Missing of 'in-text' citation and or in the list was also noticed. Inconsistencies in style format, typography and chapter organization were common lapses found. There were also evidences of insufficient proof reading in some of the reports.

Many of the research topics were mere duplication of others with no signs for new thinking to show some creativity or innovation.

Discussion

The important purposes of supervised academic research for writing a thesis or dissertation are to provide an experience in scientific writing, to familiarize the student with the steps of research, to promote critical and analytical thinking and to contribute to knowledge. The other learning objectives are to develop a skill to find out the literature, to demonstrate the appropriate depth and breadth of knowledge in the discipline, to develop the skill to evaluate the available information and thus to help them to adopt evidence based practice.

The reflected overall marginal performance in majority (Fig. 1) was due to poor scoring in more than one important section (Table 1). These added the evidences of poor achievement of the educational objectives. In Bangladesh no evaluation report was found either on the research as a whole or on the academic and supervised research. One report from Middle East, shown that the quality of research was increasing with the increased number of research work but the present investigation failed to agree with such trends. The close clustering of mean median and mode of the overall marks (Table 2) in present study were not only indicating similarity of practice but also speak in favor of common mistakes.

Evaluation on achievement of major educational objectives

Experiences in scientific writing

Very high percentages of marginal quality in almost all sections are providing evidences that in achieving this objective our learners are away from the expected level. Inappropriate chapter organization, inconsistent formatting and typography indicate that little attention was given to improve overall quality. In the present series the title was less often expository in a good number of cases as in the following-

“This clinical evaluation of non-traumatic emergency laparotomy a study of 200 cases”

This title expresses the idea of evaluating the status of patients undergoing emergency laparotomy for non-traumatic conditions but it was not clear whether the study was done for evaluation of preoperative, per-operative or postoperative condition or a compilation of all of the status. After going through the text it was found that the researcher actually described the distributions of preoperative clinical presentations and it was not an evaluation in anyway.

Similarly any reader would be in a dilemma on the researcher's intention due to the absence or undisclosed research question in a good number of abstracts. Failure to mention the methodological outline in the abstract and other weaknesses indicate that the candidates were unacquainted with the formalities of writing an abstract.

It is said that the communicating massage may be confronting due to writing style this was true in some of the reports in present study. Some are making a conclusion by quoting from others as seen in the following-

“----- (---) pointed out that the proper antenatal checkup led to reduction of feto-maternal mortality and morbidity”

Overenthusiastic deduction from a single centered cross sectional study was also there-

“The significant number of malignant case indicates global rise in malignancy”

Others' only were repeating informations from former sections in a summary form-

“The study has been carried out by------”, ‘The numbers of cases were ---’, On attempt to identify the incidences it was found that ------”, “Outcome of different processes were analyzed-----”, “There were a lot of limitations like ----”.

All these would convince anybody to believe that, having experiences in scientific writing by the learner was hardly succeeded and there is poor development in scientific and logical communication skill.

Evaluative report on the achievement of this primary objective was not found in any form after extensive search. Harold noted that the 'poor grammar and imprecise communication made the recent scientific publication standard an all-time low'. Excessive use of noun and adjectives, use of unnecessary and ambiguous words, use of complicated phrases, poor punctuation and long complicated sentences are few other weaknesses.
Absence of variable list in some cases indicates learners' confusion about the means of measuring the objectives. Boot and Beily\(^{23}\) noted that the information provided in literature review were hardly relevant to the topic in many occasions which was also found in present investigation. Presence of irrelevant information and low quality in literature review indicate that, little attention was given to this essential step. A good methodology is important for a good quality of data.\(^{24}\) The overall qualities of the methodology section in this study reflect that the importance of this section was taken lightly. Designs often did not match or was not appropriate for the title and objectives, in majority of the cases. As for example a title was written as- "Difficulty and outcome in the management of obstructive jaundice undergoing surgery: a study of fifty cases"

Where the general objective was- “To find out the relative incidence of the causes of obstructive jaundice”

And the study design was said as a- “Retrospective and prospective analytic study”

The conflict between the title and the objective and improper phrasing in both made the intention ambiguous. Likewise some other report the design was said “cross sectional prospective”. Some investigators also labeled their “case control study” as a prospective one. Completion of the study with a sample size on assumptions or on a smaller sample than that of calculated size without explaining the logic and limitation are indicative of the researchers' poor knowledge on sample size estimation and unclear concept regarding the importance of sample size. Noordzij et al.\(^{25}\) identified that this was one of the most frequent weaknesses in research. The findings of defects in methodology, title quality, abstract and other units, indicates an incomplete fulfillment of this objectives among the academic researchers under study.

**Promoting critical and analytical thinking**

One of the most important educational objectives in writing a thesis or dissertation is to promote critical and analytical thinking.\(^{5, 6}\) The low scorings in the sections like introduction, literature review, discussion, conclusion and recommendation provide an allusion to believe that critical and analytical thinking in research students are not developing as it was intended. Dietz et al.\(^{26}\) pointed out that students' generally do not work independently and neither taking part in creating new hypothesis nor involved in scientific problem solving so they do not experience how to have a critical look at research results and methods. Some researcher found that the current graduates cannot adequately solve problems and think critically in their practice.\(^{27}\) A good medical practitioner requires the ability to think scientifically and to evaluate knowledge and technology critically to solve problems.\(^{28}\) Poor formulation of research question/hypothesis and objectives and the absence of operational definitions found in the present study reflect the candidates' poverty in conceptual development. Marginal quality and mere repetition of results in many of discussions in this series indicates researchers' failure to conceive the purpose of writing a discussion. This was also indicative of poor comprehension of their own result to apprehend the specific objectives.

It is said that the dissertation discussion part should serve as the interpretation of the hypothesis or should answer the research question.\(^{5, 6, 29}\) In this series, answer to the research question and evidences supportive of hypothesis was not clearly correlated with 'intelligent agreement of findings' in majority of the reports. All these findings lead to the suspicion on the success in promoting the critical thinking and evaluation skill among the learners.

**Contribution to knowledge**

In present series almost all had simply repeated the investigation done by others without a reason and at times just replicated the whole process without any justifiable argument or a new context for repetition. There was hardly any evidence of innovation and creativity to justify the repeat work. The high rate of replicative works and very low overall quality indicate their trivial contribution in knowledge base. It was found in a study that only less than 3% of the published articles in academic journals are useful.\(^{30}\) Medical dissertation/thesis also contributes very little to knowledge base, because of their low quality and very few (17 %23.8%) of them are published in journals for a wider dissemination.\(^{31, 32}\) In Bangladesh there was no known data available on this issue but it could be presumed very low.

The present study added the evidences in favor of the assumption that the quality of thesis/dissertation in the medical institutes is away from expectation. The basic quality as a scientific writing is hardly ensured. There were also signs of inadequate supervision and gaps in the examination process. From an overview of the finding one could have doubts about the achievement of the educational objectives from such a time and resource consuming academic activity.

To address this issue planning, designing, supervision and evaluation system of supervised research in relation to the postgraduate medical qualification of Bangladesh needs critical attention. Attention should be given on further research to focus and find out the problems and their solution in the supervised academic research. To determine the effective system of supervision and to enforce more effective system of examination, educational researchers should come forward.

**Limitations and scopes of the study**

The major limitations in this study were the sampling technique and sample size. For the purpose of generalization it would be better if the study would be carried out by taking
data from all or most of the institute throughout the country. The rating scale and the rubrics may be another potential area of weakness, instead of repeated pretesting; as they were freshly prepared by the present researcher, they may have to be improved further. Inspite of all limitations, as a beginning of this type of study in this country and most probably the first one in evaluating the thesis/dissertation in higher educational level, the information obtained might be a useful foundation for future works.

References


3. WFME. Postgraduate Medical Education, WFME Global Standards for Quality Improvement, World Federation for Medical Education (WFME), Copenhagen 2003; http://www.wfme.org


10. Tyler RW 1949, How can the effectiveness of learning experience be evaluated in Basic principles of curriculum and instruction; The University of Chicago Press; Chicago, pp. 104-23.


