

EMPIRICAL RESEARCH

Open Access



Evaluation of immediate impact of Faculty Development Programme using a pretest–post-test study design format

Shahid Hassan , Sunil Pazhayanur Venkateswaran and Vishna Devi Nadarajah

Abstract

Background: Workshops are the most common models to enhance knowledge and skills in a specific subject area with an intent to explore, solve a problem and/or innovate new things. The most important aspect of a workshop is the transfer of knowledge in a safe learning environment as a faculty development activity (FDA). At International Medical University (IMU), Malaysia's first private medical university which was established in 1992, Faculty Development Programmes (FDPs) are run throughout the year in order to enhance the knowledge and skills in teaching and assessment. In order to sustain this faculty development, IMU has a dedicated medical education unit called the IMU Centre of Education (ICE) with dedicated staff and respected faculty developers who are academic role models to the faculty of the institution. However, FDA are collaboratively run by ICE and IMU Centre for lifelong learning (ICL).

Objectives: To determine the immediate impact of faculty development workshops for health professionals in teaching schools of IMU to enhance the teaching and assessment abilities of the faculty.

Methodology: A retrospective quantitative research design was developed to collect data from multiple standard setting workshops using a 3-point Likert scale. A 20 items questionnaire as a pretest from the participants with and without the prior reading of online posted reading materials. An interventional hands-on workshop and a post-test score, using the same 20 items questionnaire, followed the workshop intervention. A collated quantitative data were gathered from a sample of 139 participants attending the standard setting workshops. Data were analysed using paired *t* test, one-way ANOVA and ANCOVA with effect size in SPSS version 24.

Results: A mean difference between pretest and post-test score was significant at $t(138) = 92.24, p < 0.01$. A barely significant difference of mean scores between pre-read, partially read and not-read participants was found at $F(1, 136) = 9.402, p = < 0.05$ and $\eta^2 = 0.121$ by one-way ANOVA. A post-test difference of the mean scores across those read, partially read and not-read for reading materials on a controlled pretest score determined by one-way ANCOVA remained non-significant at $F(1, 136) = 0.240, p = 0.787$ associated with a practical effect size = 0.4% only.

Conclusion: A significant difference of the mean pretest and post-test score within the group was also significant between the groups. A post-test score, controlling on pretest score, was found not significant and is suggestive of an effectively delivered workshop for all participants. As a practical guide, a 7-step plan is recommended to ensure immediate impact of FDP. Majority of the participants rated the workshop as good in terms of content, organisation,

*Correspondence: shahidhassan@imu.edu.my
School of Medicine and IMU Centre for Education, International Medical University (IMU), 126, Jalan Jalil Perkasa 19, Bukit Jalil, 57000 Kuala Lumpur, Federal Territory of Kuala Lumpur, Malaysia



© The Author(s) 2021. This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

delivery and usefulness. A high percentage of survey respondents reported that similar workshops to be offered in future.

Keywords: Faculty Development Programme, Workshops, Pretest, Post-test, Quantitative evaluation

Introduction

Faculty development is important for inculcating meta-cognitive awareness in teachers about what they know and what they need to know. Faculty Development Programme (FDP) in this context is therefore a necessity for institutions to build their faculty motivation and confidence and to help ensure that the educational goals are met with as per the vision of an institution. Professional organisational experts advocate greater awareness about the acquisition of knowledge in teaching and learning through comprehensive faculty development (General Medical Council, 2003; Levinson, 1998). FDP plays a significant role in undertaking on the job learning, beneficial to the career of the faculty (Jehanzeb & Bashir, 2012).

FDPs are also an important part of health professions education and in delivering the medical curriculum efficiently (Steinert et al., 2016; World Health Organization, 2013). Besides teaching, the role of the faculty is equally important in assessment for successful implementation of curriculum in medical education. Assessment developed for learning helps in many ways, from taking logical decisions on pass/fail, to providing feedback to students and insight to teachers in order to plan and guide instructional strategies. Standard setting of assessment done through a blend of judges' decision, psychometrics and practicality provides informed judgment for assessment to be defensible, credible and supported by literature (Barman, 2008). To define the minimally competent borderline students and the standards to make logical decisions on pass and fail in an undergraduate programme, a faculty development training in standard setting is considered essential for health professional educators engaged in assessment. The present study involves the efficiency of such a training programme delivered as repeatedly conducted workshops focused on basic key concepts and challenges in standard setting. The effort was to improve the quality of and transparency of assessment with regards to standard setting practised both for written and clinical examinations. The intended outcome is an effective faculty development, which is more credible and defensible for decision makers involved in standard setting for assessment practised in International Medical University (IMU).

While planning and implementing the FDP, it is important to consider the environment in which teaching and learning will take place and where the actual curriculum will be delivered (Koens et al., 2005). Whereas Faculty

Development Programme (FDP) has become a need for teachers engaged in health professions education, it is equally important to evaluate the impact of training for its effectiveness (Lancaster et al., 2014). Higher education demands to improve faculty development across all levels of faculty experience, especially in the area of teaching and learning inclusive of assessment (Cook & Kaplan, 2011). Gillespie and colleagues recommend 10 steps to be considered when building a Faculty Development Programme with its long-term impact (Stes et al., 2007), which also includes development and guideline principles, clear goals, and assessment procedures. In the present study, the authors' concern has been the assessment of FDP to produce evidences to continuing faculty enhancement in teaching and learning skills of faculty in medical education.

Understanding the factors associated with the students' poor performance and their real score minus errors, the students' ability to learn may not be the only reason. The other important reason could be the teachers' ability to facilitate and assess (Steinert et al., 2006). In medical schools, often the recruitments are based on faculty's expertise in their subject specialty and not on their ability to teach and assess. Medical schools therefore put emphasis on faculty development in teaching and assessment for them to become effective supervisors. Assessment using pretest and post-test methods with in-between intervention strategies if used in training workshops may bring about positive changes in the teacher's attitude, knowledge and skills (Ramalingaswami, 1989). Pretest and post-test continuous scores can easily be analysed for significant outcomes by using paired *t* test for knowledge and skills acquired during the workshop by the participants with diverse learning styles and education backgrounds (Mokkapati & Mada, 2018).

Methodology

The Standard Setting Workshop is conducted once or twice a year at our institution and is delivered at 2 levels (Basic/Intermediate and Advanced). The 8 workshops were delivered face to face and conducted between the years 2017 and 2019 at IMU. Out of a total of 146 faculty and staff attending these 8 workshops, 139 were included in the present study meeting the inclusion criteria (Table 1).

Participants were self-assigned to three subgroups based on their preclass reading of circulated reading

Table 1 Standard setting workshop offered as Faculty Development Programme in IMU, Malaysia

No	Date, year and (hours) of workshop	Number of faculty and staff registered	Number of faculty and staff attended
1	2 May 2017 (9 h)	36	34
2	4 May 2017 (8 h)	16	15
3	9 November 2017 (8.5 h)	12	10
4	7 June 2018 (8 h)	40	38
5	19 September 2018 (8.5 h)	15	10
6	12 December 2018 (9 h)	12	11
7	18 September 2019 (8.5 h)	10	10
8	21 May 2019 (5 h)	22	18
	Total	163	146 (Included in study 139)

materials indicating their preparedness to attend the workshop. A Google-form survey immediately after their response to the pretest was administered to define the three levels of independent variable of participants.

The learning outcomes of the workshop were as follows:

1. Define the common standard setting methods practised in medical education.
2. Perform the Modified Angoff procedure for a written test of One Best Answer (OBA) MCQ.
3. Perform the Borderline Regression Method (BRM) to determine the cut-off score in a clinical test of OSCE.

The prerequisites to attending the workshop for the participants included reading the materials sent prior to attending the workshops, responding to the pretest and post-test questions, preparing 2–3 OBA items from their area of expertise and sharing these questions during the group work activities. It was also required for participants to have a knowledge of Microsoft Excel™ to understand the hands-on sessions better.

The study employed a quantitative research design with descriptive and comparative analysis of pretest and post-test scores using convenient sampling to estimate the causal impact and efficiency of the intervention on targeted participants without randomly assigned groups. The intervention was a one-day workshop with few introductory interactive lectures and hands-on sessions on standard setting in student's assessment for determining the cut-off score of examination questions.

Data from the standard setting workshops were collected on a 3-point Likert scale of 20 items questionnaire

Table 2 Change in pretest and post-test score before and after interventional instructions ($n = 139$)

Variable pretest and post-test scores	Pretest score mean (SD)	Post-test score mean (SD)	Mean score difference (95% CI)	<i>t</i> -statistic (<i>df</i>)	<i>p</i> value
	9.96 (0.974)	18.46 (1.205)	− 8.5 (− 8.67, − 8.31)	92.24 (138)	< 0.001

(see “Appendix”) as a pretest determined with and without the prior reading of circulated reading materials, followed by a one-day interventional training and a post-test using the same questionnaire. A sample size using the standard error of mean formula was employed at 95% confidence interval and a pilot study to estimate 139 subjects attending the standard setting workshops from May 2017 to May 2019.

Results

Quantitative component

The pretest and post-test scores of 139 out of 146 participants based on the inclusion/exclusion criteria were analysed (see Table 1). For differences within the group, we ran a paired *t* test and for the differences between groups we ran a one-way ANOVA. Post hoc test was further performed to determine at least one difference between groups. Controlling for pretest covariance effect between group ANCOVA was also conducted after meeting the assumptions of equality of regression coefficient as well as equality of homogeneity of variance across the variables. A mean difference of 8.5 between the pretest score = 9.96 (0.974) and post-test score = 18.46 (1.205), respectively, was significant at $t(138) = 92.24$, $p < 0.01$ (see Table 2) associated with a high correlation of 0.520.

A one-way ANOVA result in this study to evaluate the null hypothesis is that there is no difference of post-test scores among the participants based on their pre-class reading of reading material circulated prior to the workshop. A Google-form survey was undertaken together with the pretest to establish the participants' groups read vs partially read or not read any of the reading material. Majority of participants accepted the reading material was not read though (see Table 3) and showed three groups: read (mean = 19.03, SD = 0.873, $n = 39$), partially read (mean = 18.53, SD = 1.014, $n = 45$) and not read (mean = 18.00, SD = 1.376, $n = 55$).

The assumption of homogeneity of variance test using Leven's test was violated since significant at $F(2, 136) = 3.865$, $p = 0.023$. The alternative Welch statistics as a robust test of ANOVA was found significant $F(2, 136) = 9.402$, $p = 0.000$, $\eta^2 = 0.121$. Thus, a significant

difference of participants' performance is shown in the post-test based on their pre-class reading status. However, the actual difference between groups was quite moderate, based on Cohen's (1988) convention for interpreting effect size, $\eta^2 = 0.121$ and partial eta squared = 12.1%, which is a low effect size for Cohen's *d* test as well as for partial eta squared that suggests only 12% of variation in post-test score and can be explained by a change in pretest score. Subsequently based on the ANOVA result of a significant difference of pre-class reading status, a post hoc comparison to evaluate pairwise difference among groups means was conducted (see Table 4). This was done using a robust test of Games–Howell instead of Tukey's test for assumption of homogeneity of variance violated. The test revealed a significant difference between the mean score of those who read the reading material and did not read the reading material, $p = < 0.05$. Participants in the partially read preclass status did not significantly differ from the two groups, $p = > 0.05$.

ANCOVA on the other hand after having met the assumptions of regression of coefficient, $F = 0.541$, $p = 0.583$, showed no significant difference, $F = 0.240$,

$p = 0.787$, among the three preclass groups based on their preparedness for the workshop by their status of prior reading (see Table 5). However, a substantially low partial eta squared = 0.004 (0.4%) was seen. No significant effect of pretest status is also shown by the adjusted mean score uncontaminated by the effect of covariant pretest as follows, mean of read group = 18.413, mean of partially read group = 18.385 and mean of not read group = 18.556. Looking at the interaction of pretest times and preclass groups has also shown no significant difference, $F = 0.541$, $p = 0.583$; however, a slightly higher partial eta squared = 0.8%.

Discussion

The academic service to the institution is linked with the professional development of its faculty members. Faculty Development Programme (FDP) also ensures educator's greater awareness and acquisition of knowledge in teaching and learning through comprehensive professional training. FDP, popular as Faculty Development Activity (FDA) in IMU, is the integral training activity for medical teachers to learn about teaching, learning and assessment in order to improve the quality of medical education both

Table 3 Post-test score before and after interventional instructions ($n = 139$) by one-way ANOVA

Variable prior reading	Number	Mean (SD)	<i>F</i> -statistic (<i>df</i>)	Partial eta squared	<i>p</i> value
Fully read	39	19.03 (0.873)	9.402 (2, 136)	0.121 (12.1%)	> 0.05
Partially read	45	18.53 (1.014)			
Not-read	55	18.00 (1.376)			

Table 4 One-way ANOVA result for multiple comparison post hoc test (between the group)

(I) Assessment mode group	(J) Assessment mode group	(I–J) mean difference	<i>p</i> value	95% CI	
				Lower bond	Upper bond
1 Preclass read	2	0.492	0.122	– 0.10	1.08
	3	1.026*	0.000	0.46	1.59
2 Preclass partially read	1	– 0.492	0.122	– 1.08	0.10
	3	0.533	0.055	– 0.01	1.08
3 Preclass not read	1	– 1.026*	0.000	– 1.59	– 0.46
	2	– 0.533	0.055	– 1.08	0.01

Table 5 Post-test mean score across those read, partially read and not read the reading materials on a controlled pretest covariance by one-way ANCOVA

Variable post-test score	Number	Mean score (SD)	<i>F</i> -statistic (<i>df</i>)	Partial eta squared	<i>p</i> value
Fully read	39	19.03 (0.873)	0.240 (1, 136)	0.004 (0.4%)	> 0.05
Partially read	45	18.53 (1.014)			
Not-read	55	18.00 (1.376)			

professionally and personally and in keeping with global guidelines (McLean et al., 2008). Training acquired in a systematic manner helps to develop the knowledge, skills required by an individual to perform adequately well on a given task (Guglielmo et al., 2011).

The current study is based on a comprehensive faculty development plan focused on outcome and evaluation for immediate impact on which there are limited articles in the literature (Steinert et al., 2006). Emphasis on the current FDA was on subjects selected for FDP to be contextualised and relevant to the faculty and institution's need assessment. Every FDP consumes time, efforts and cost and therefore needs to be assessed for its benefit not only for faculty but also for institutions (Armstrong & Barsion, 2006). The quantitative data analysis supported the overall positive impact and efficiency of the workshops.

Majority of the participants in the current study admitted that they have not read the circulated articles, which was reflected in their comparatively low score in the pretest. However, an unexpected higher difference of pretest and post-test mean scores was observed among those not read vs. partially or not read participants as shown in the estimated marginal means plot (see Fig. 1). This might have been due to extra efforts applied during the workshop by the attending participants without preparation for the workshop. However, difference of means scores was highly significant ($8.5, p = < 0.001$) within the group (see Table 2) in pretest (mean = 9.96, SD = 0.976) and post-test (mean = 18.46, SD = 0.1.205) before and after the intervention, respectively, using paired *t* test and is compatible with the literature (Baral et al., 2012; Dhungana et al., 2015). An overall improvement in the faculty knowledge after the workshop as shown by the improvement in the post-test scores over the pretest scores indicates an immediate positive impact of the workshop.

Long-term impact is difficult to gauge for training workshops. However, an immediate impact can be developed and designed to measure the FDP for statistical as well as practical evidences of its effectiveness. In the current study, a comprehensive method based on statistical tests for immediate impact has been suggested to establish evidences of its effectiveness. The analysis of immediate impact can be viewed with other steps of setting up of learning outcomes, identifying targeted audience to benefit, calculating the time, efforts and costs. For an immediate impact, a pretest and post-test strategy as in the current study was incorporated with or without pre-class activities such as assignments or reading of identified articles or text sent out prior to attending the training workshops. Data were collected from a well-written pretest/post-test questionnaire from the subject area and the participants' response to determine three self-assigned groups using a survey. The three groups determined are

read, partially read or not read status of preclass reading of circulated materials prior to attending the training workshops. Data collected were analysed using statistical tests including paired *t*-test, one-way ANOVA with effect size and one-way ANCOVA with effect size for drawing conclusions (see Fig. 2).

On the other hand, a significant difference between groups in one-way ANOVA (see Table 3) suggests that reading was comprehensively and effectively done to apply new knowledge to select the correct answers in the pretest. However, a post hoc ANOVA revealed the difference in mean scores was between group 1 (read) and group 3 (not read). There was no significant difference between the read versus partially read and partially read versus not read groups (see Table 4). Another interesting finding was on the post-test score controlling pretest knowledge, which was found non-significant using one-way ANCOVA (see Table 5). This confirmed the earlier finding of reading through those circulated articles was not comprehensive and effective enough. However, post-test score controlling on pretest was insignificant and is another indicator of an effectively delivered workshop for all participants irrespective of their prior preparedness status for a training session. In FDP looking for immediate impact, the authors suggest that the evaluation should begin with writing of the workshop proposal including the following steps (see Table 6): (1) defining the learning outcomes of training program. (2) Identifying the faculty, who may benefit from faculty development training. (3) Selecting facilitators with their expertise to satisfy audience gaps in knowledge. (4) Calculating the exact time, efforts and cost required for delivering the workshop within the resources available. (5) Identifying the learning resources and teaching materials for outside class activity. (6) Administering a pretest comprising of 15–25 items questionnaire from the relevant subject area and the same to be used as post-test. (7) Evaluating for immediate impact using a battery of statistical tests to analyse the data collected as pretest and post-test continuous scores.

A clear plan to evaluate the impact of workshop is always needed. Ultimately it is of paramount importance to measure the impact of workshop in the first place and balance it with the amount of time and effort required to organise and evaluate the workshop. A significant difference of mean pretest and post-test score within groups was not the same between the groups. A post-test score, controlling on pretest score found not significant, suggests an effectively delivered workshop for all participants irrespective of their preclass reading. Majority of the participants rated the workshop as good in terms of content, organisation, delivery and usefulness in their follow through feedback in routine. A high percentage of

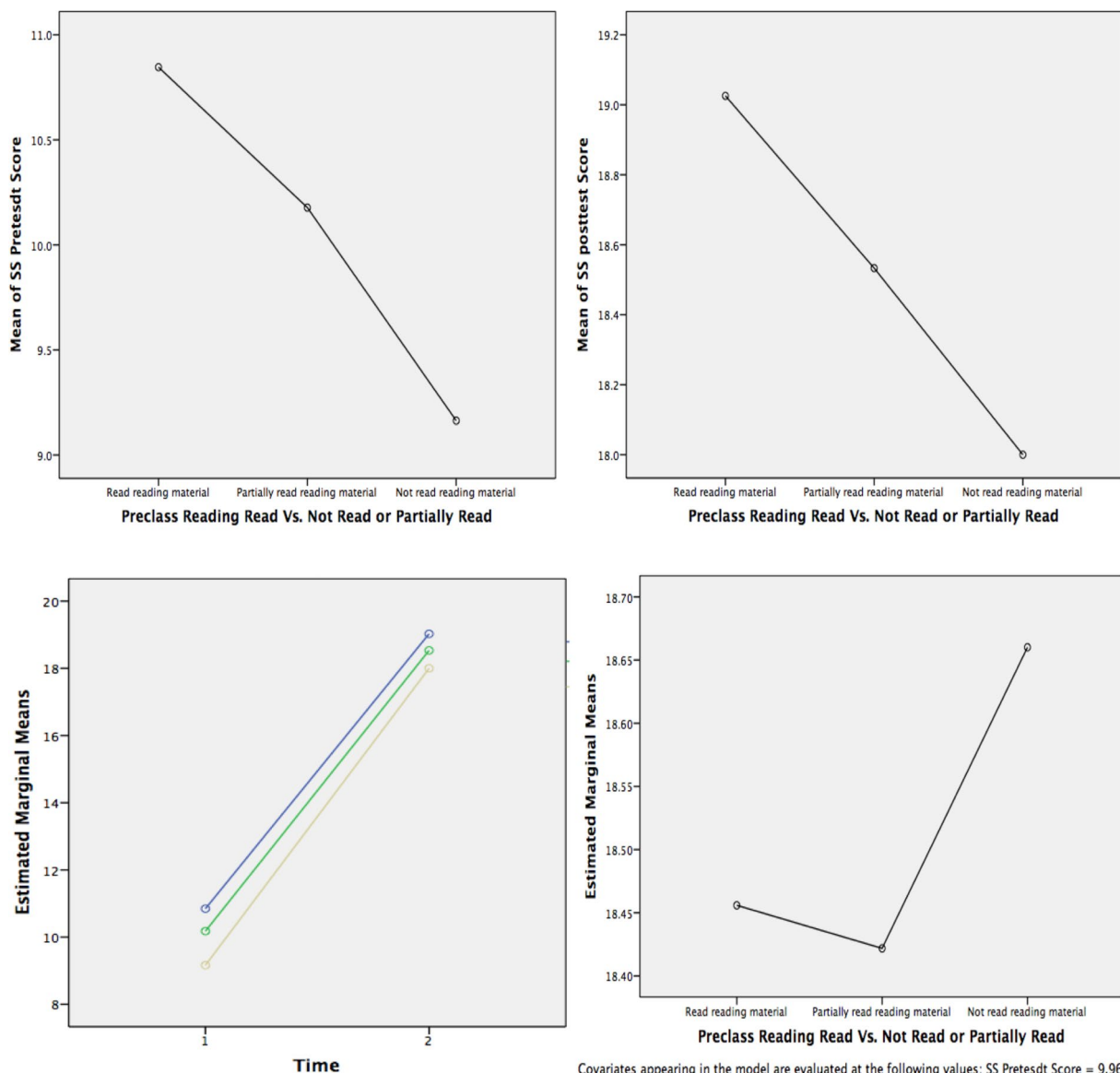


Fig. 1 A significant but small mean difference in pretest (Plot 1) and post-test score (Plot 2 and 3). However, an insignificant difference is seen after controlling the pretest covariance effect (Plot 4) with a change order of distribution among the preclass status of reading materials (not read is shown higher mean than the other two groups)

respondents wanted similar workshops to be offered in future. However, a feedback survey tailored to the needs assessment of individual faculty development activity (workshop) be designed and developed to receiving the meaningful feedback from the participants for qualitative analysis.

Findings from the current study suggest that a comprehensive plan inclusive of evidences collated on immediate impact using a quantitative method is

imperative to the effectiveness of an FDA. Strategies must be set to send out relevant reading materials, brainstorming participants on their prior knowledge with a flipped classroom approach, pretest and post-test using comprehensive analytic statistical methods (practised in this study). Feedback often sought from the participants is a generic routine approach for all FDAs; however, a feedback survey tailored to the needs assessment of individual faculty development activity (workshop) is recommended.

Statistical methods for Quantitative Analysis to the Determine Immediate Impact of FDP

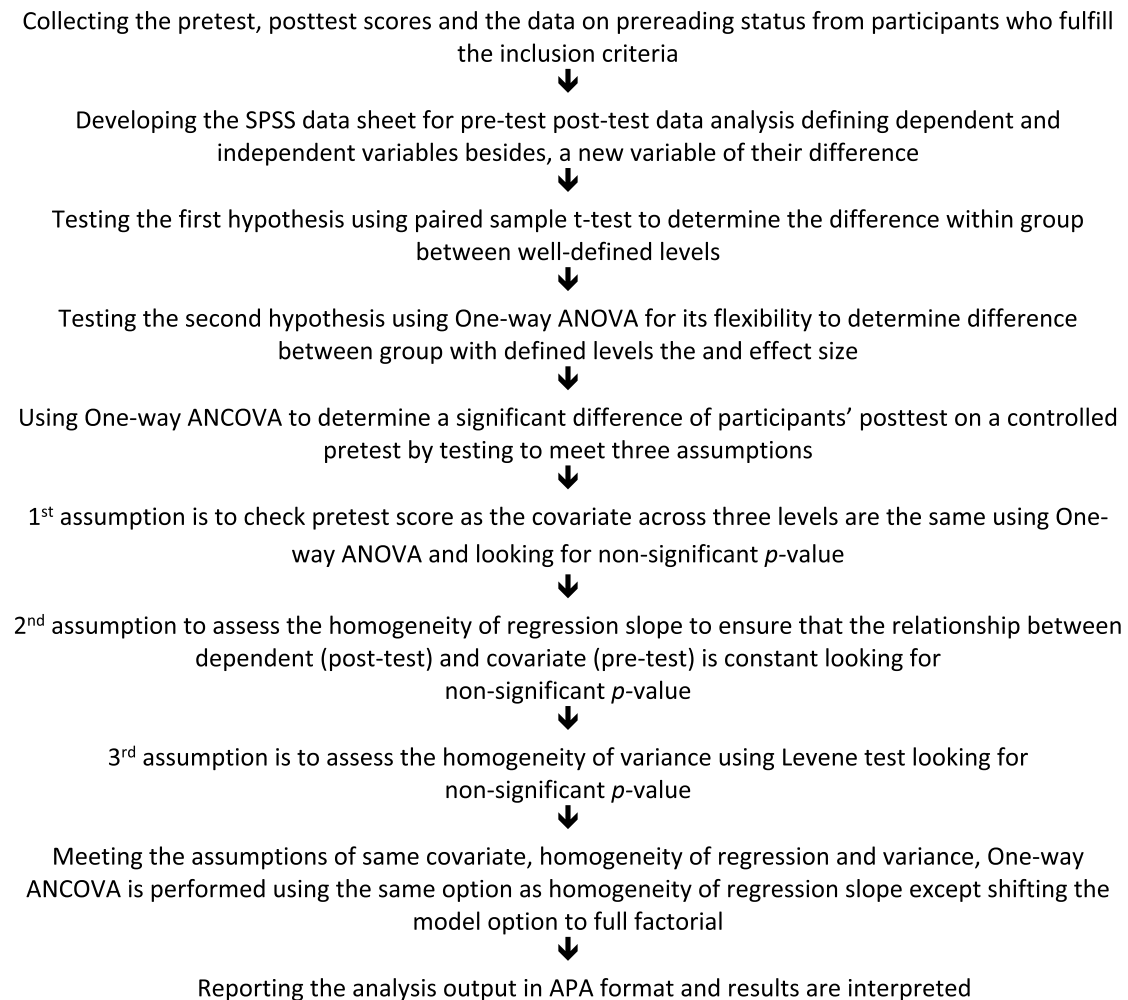


Fig. 2 Flow chart of statistical methods to measure an immediate impact of faculty development activity

Table 6 Practical guide and steps to measure immediate impact of Faculty Development Programme

Sequential steps to measure the immediate impact of Faculty Development Programme

1. Defining the learning outcome of a single titled training program
2. Identifying the faculty, who may benefit from Faculty Development Programme
3. Selecting facilitators with their expertise to satisfy audience gaps in knowledge
4. Calculating the exact time, efforts and cost required to deliver within the resources available
5. Identifying learning resources and teaching materials for outside class activity
6. Writing and administering a 15–25 items questionnaire to test the audiences' prior knowledge as pretest and post interventional knowledge as post-test
7. Evaluating the pretest and post-test scores for immediate impact using statistical method to test the hypothesis developed in a pretest and post-test research design

Conclusion

FDP may be an effective predictor of curriculum and educational reforms. Analysis of FDP for immediate impact is considered essential for effectiveness of on the job faculty training and faculty role as a change-agent is required for innovative curriculum review. On the job faculty development is a good practice to help health professions educators in updating their knowledge and skills required by the institution and accreditation bodies such as Malaysian Qualifications Agency as well as other international and regional professional bodies as an evidence of progressive curriculum development. Evaluation of such programs with positive evidences of immediate impact may strengthen the institutional efforts of effectively implementing the curriculum.

Limitations and recommendations

The FDA for its immediate impact was evaluated using a quantitative study design. However, in future all FDA workshops should use a mixed-method approach (Focus group discussions), drawing on both, quantitative and qualitative data analysis. The mixed methods are useful in formulating a deeper understanding of the research problem, explaining quantitative results with a qualitative follow-up and analysis. A mixed methods design for its qualitative data may assist to explain in detail the initial quantitative results (Creswell & Plano Clark, 2011).

Appendix: Standard setting method for examination questions: pretest/post-test

Name: _____ Discipline: _____
 Institution: _____ Date: ____/____/____.
 About the reading material: 1. I did not read _____
 2. I partially read _____ 3. I fully read all _____

Instruction: Please read each item on standard setting carefully and tick (✓) the appropriate column if right.

For agree and disagree correctly answered = 1 mark, incorrectly answered = -1 mark and Not sure = 0 mark
 Test Outcome: 15–20 (Excellent), 14–17 (Very Good), 11–13 (Satisfactory) and ≤ 10 (unsatisfactory).

No	Items	Agree	Not sure	Disagree
1	Standard setting method is the same as standard setting strategy that logically decides a test outcome			
2	Relative standards are most appropriate for examinations where the purpose is to identify a certain number of examinees			
3	Standard setting is based on expert judgment, informed by data related to a specific cohort performance			

No	Items	Agree	Not sure	Disagree
4	A standard is a gathering of professional values judgments through the majority expression on a single score on a test			
5	Standard setting is a technique for divining a scientifically and absolutely correct solution			
6	Absolute standards are most appropriate for tests of competence			
7	Fixed Percentage Method is standard setting independent of test content and how much of it the examinees know about content			
8	In Angoff's method, judges produce absolute standards, which is best suited to tests that seek to establish competence			
9	In Angoff's method, judges are informed by administrators of the characteristics of a borderline group of examinees			
10	In Angoff's method, judges consider difficulty and importance of items on test as key factor to know student can correctly answer			
11	In contrasting group method, the scores of the selected passers and failures are graphed and the cutting score is calculated as the point of least overlap between the distributions			
12	In "Contrasting Groups Method", judges are directly informed by the actual test performances of the examinees			
13	In Hofstee method, it allows examiners' direct manipulation of the false positive and false negative rates			
14	In a high stakes examination six to eight judges are a minimum for the standard setting method			
15	With minimum number of judges and one or two outliers, it may be reasonable to use the median instead of the mean			
16	Reliability by Cronbach's alpha is considered an absolute method for determining the unidimensionality of test items			
17	A low inter-item correlation mean average is known to affect coefficient alpha in the presence of increase number of items			
18	Modified Angoff is a recommended method for standard setting of objective structured clinical examination			
19	Borderline Regression Method (BRM) is a time consuming labour-intense process to decide on cut-point-score on OSCE			

No	Items	Agree	Not sure	Disagree
20	Root Mean Square Error (RMSE) is a better option than Cronbach's alpha to evaluate the reliability of OSCE			

Acknowledgements

Authors would like to acknowledge IMU-JC for approval of the research project and allowing funds as requested, ICE (IMU Centre for Education) and ICL (IMU Centre for Lifelong Learning) for organising the FDA in IMU and eLearning Department.

Authors' contributions

SH is the principle researcher (PI) in this research project, who has written the proposal together with the other two co-researchers and defended the proposal in the IMU Joint Committee for Research and Ethics (IMU-JC). SH has also written the manuscript, which was reviewed by the SPV. SH is also the corresponding author. SPV has contributed in defending the project together with other two authors and has also been engaged in writing and reviewing of the manuscript. VDN has played a scholarly role by forwarding the idea of researching the immediate impact of faculty development activities in IMU and will be engaged in addressing the peer reviewers comment. All authors read and approved the final manuscript.

Funding

The research proposal was presented in 195th IMU Joint Committee on Research and Ethics on 13 February 2020 and was given full approval with reference No. 4.20/JCM-195/2020 and Project ID no: IMU 471/2020.

Availability of data and material

Data included in the manuscript are readily available for the journal publication process.

Declarations

Competing interests

No conflict with anyone at individual or institutional level with regard to content and material is used in this manuscript. The authors declare that they have no competing interests.

Received: 26 May 2020 Accepted: 31 March 2021

Published online: 21 April 2021

References

- Armstrong, E. G., & Barsion, S. J. (2006). Using an outcomes-logic-model approach to evaluate a faculty development program for medical educators. *Academic Medicine*, 81(5), 483–488. <https://doi.org/10.1097/01.acm.000022259.62890.71>.
- Baral, N., Gautam, A., Lamsal, M., Paudel, B. H., Das, B. K. L., & Aryal, M. (2012). Effectiveness of teachers' training in assessment techniques: Participants' perception. *Kathmandu University Medical Journal*, 9(3), 189–192. <https://doi.org/10.3126/kumj.v9i3.6303>.
- Barman, A. (2008). Standard setting in student assessment: Is a defensible method yet to come? *Annals of the Academy of Medicine, Singapore*, 37, 957–963.
- Cook, C. E., & Kaplan, M. (2011). *Advancing the culture of teaching on campus: How a teaching center can make a difference*. Stylus Pub.
- Creswell, J. W., & Plano Clark, V. L. (2011). *Designing and conducting mixed methods research*. (2nd ed.). Sage Publications Inc.
- Dhungana, G. P., Piryani, R. M., Chapagain, M. L., & Neupane, M. (2015). Effectiveness of teacher training conducted at Chitwan Medical College,

- Bharatpur, Nepal. *Journal of Chitwan Medical College*, 5(2), 1–5. <https://doi.org/10.3126/jcmc.v5i2.13147>.
- General Medical Council. (2003). *Tomorrow's Doctors. Report on Education Committee of the GMC*. General Medical Council, London.
- Guglielmo, B. J., Edwards, D. J., Franks, A. S., Naughton, C. A., Schonder, K. S., Stamm, P. L., Thornton, P., & Popovich, N. G. (2011). A Critical appraisal of and recommendations for faculty development. *American Journal of Pharmaceutical Education*, 75(6), 122. <https://doi.org/10.5688/ajpe756122>.
- Jehanzeb, K., & Bashir, N. A. (2012). Training and development program and its benefits to employees and organizations: A conceptual study. *Far East Journal of Psychology and Business*, 9(2), 58–71.
- Koens, F., Mann, K. V., Custers, E. J., & Ten Cate, O. T. (2005). Analyzing the concept of context in medical education. *Medical Education*, 39, 1243–1249. <https://doi.org/10.1111/j.1365-2929.2005.02338>.
- Lancaster, J. W., Stein, S. M., Maclean, L. G., Amburgh, J. V., & Persky, A. M. (2014). Faculty Development Program models to advance teaching and learning within health science programs. *American Journal of Pharmaceutical Education*, 78(5), 99. <https://doi.org/10.5688/ajpe78599>.
- Levinson, W. (1998). Clinician-educators in Academic Medical Centers: A two-part challenge. *Annals of Internal Medicine*, 129(1), 59. <https://doi.org/10.7326/0003-4819-129-1-199807010-00013>.
- McLean, M., Cilliers, F., & Van Wyk, J. M. (2008). AMEE Guide No. 36: Faculty development: Yesterday, today and tomorrow. *Medical Teacher*, 30, 555–584.
- Mokkapati, A., & Mada, P. (2018). Effectiveness of a teacher training workshop: An interventional study. *Journal of Clinical and Diagnostic Research*. <https://doi.org/10.7860/jcdr/2018/30165.11219>.
- Ramalingaswami, V. (1989). Medical education: How is change to come about? *Medical Education*, 23(4), 328–332. <https://doi.org/10.1111/j.1365-2923.1989.tb01557.x>.
- Steinert, Y., Mann, K., Centeno, A., Dolmans, D., Spencer, J., Gelula, M., & Prideaux, D. (2006). A systematic review of faculty development initiatives designed to improve teaching effectiveness in medical education: BEME Guide No. 8. *Medical Teacher*, 28(6), 497–526. <https://doi.org/10.1080/01421590600902976>.
- Steinert, Y., Mann, K., Anderson, B., Barnett, B. M., Centeno, A., Naismith, L., Prideaux, D., Spencer, J., Tullo, E., Viggiano, T., Ward, H., & Dolmans, D. (2016). A systematic review of faculty development initiatives designed to enhance teaching effectiveness: A 10-year update: BEME Guide No. 40. *Medical Teacher*, 38(8), 769–786. <https://doi.org/10.1080/0142159x.2016.1181851>.
- Stes, A., Clement, M., & Petegem, P. V. (2007). The effectiveness of a Faculty Training Programme: Long-term and institutional impact. *International Journal for Academic Development*, 12(2), 99–109. <https://doi.org/10.1080/13601440701604898>.
- World Health Organization. (2013). *Transforming and scaling up health professional education and training: Policy brief on faculty development*. World Health Organization.

Shahid Hassan Prof. Dr. Shahid Hassan completed his MHPE from Maastricht University, Netherlands, and worked as medical educationist in USM, Head of Unit of Medical Education in UPM, UniSA, and currently as Interim Director IMU Centre for Education. As clinician in ENT his subspecialty is Head and Neck and Thyroid Surgery. He has been awarded with WIPO and ITEX Gold Medal and Excellent Scientist Award for Innovative research. He has published 115 research papers and two books. He is the Past Associate Editor of PJORL and Member Editorial Board of a number of journals. He has been the Vice President and EXCO of MAEMHS and Leader of Assessment Subgroup MedNet Malaysia.