Assessing the Effectiveness of a New Curriculum: Part I

Marie E. Dagenais, D.M.D.; Dana Hawley, D.M.D.; James P. Lund, B.D.S, Ph.D.

Abstract: Although it is important to assess the effectiveness of programs, courses, and teaching methods to ensure that goals are being achieved, it is very difficult to evaluate the impact of fundamental changes in a whole curriculum. This paper reviews measures that have been used in the past in dentistry and medicine for evaluating academic programs: curriculum guidelines; competency documents; discussion and focus groups; competency examinations; board examinations; oral comprehensive examinations; student, alumni, and patient satisfaction surveys; evaluation by instructors; and clinical productivity. We conclude that, since no standard method exists, several tools should be used to obtain a multidimensional assessment.

In 1995, the Faculty of Dentistry at McGill University adopted a new curriculum in which the basic sciences program is shared with the Faculty of Medicine. As a first stage in the assessment of this new curriculum, we reviewed the methods that have been used previously to measure the impact of major curricular changes in health sciences faculties.

Rationale

The major reason for assessing the effectiveness of a curriculum is to ensure that the goals of the program have been achieved: in the words of Casamassimo,1 “finding out how you are doing in doing what you do.” The proponents of change need to be reassured that the right decision was made and that competent dentists, well prepared to face the challenges of dental practice, are being produced. Furthermore, as Chambers2 has pointed out, there is an increased public and professional demand for accountability in education. This includes a need to prove to the students themselves, to the profession, and to the public at large that the students who graduate from the new curriculum are well equipped to address the changing oral healthcare needs of the community.

Although many educators recognize that the practice of dentistry is evolving and that the dental curriculum must adapt accordingly, many people are resistant to change.3-7 Faculty resistance was identified as the third most significant barrier to modifications of the dental school curriculum.8 This may explain why the basic predoctoral dental curriculum has not significantly changed in ninety years.9 While the content has evolved to reflect science and technological developments, the basic course structure and methods used to train students have not. After McGill adopted a new curriculum, the belief of many part-time instructors that the old curriculum was better because students did more laboratory work has had to be dealt with constantly despite scientific evidence that the duration of laboratory training is not the main determinant of clinical skills. For instance, Caminiti10 has shown that a short intensive course can be as efficient as two years of preclinical work in training students to place four implants in the maxilla. Finally, assessment of a new academic program allows for problems to be identified and solutions to be proposed.

The effectiveness of a new curriculum should be measured against a standard—which is usually the old curriculum. To run an ideal educational trial, incoming classes should be randomly divided into...
two groups: one following the old curriculum, the other the new. In the medical literature, many studies have compared the outcomes of problem-based and conventional curricula. The medical schools of the University of New Mexico and Harvard have run parallel problem-based learning and conventional curricula. Other investigators have compared the performance of students from several schools with different curricula, which introduces significant limitations on interpretation of the results.

For many dental schools including McGill, managing two equal parallel paths of instructions is impossible. Two programs could be conducted simultaneously to evaluate a course, but this is generally unrealistic for an entire curriculum. Some schools have tried to compare programs directly by placing a small cohort of students in the new program in parallel with the old curriculum. Nonrandom assignment of students, staff, and resources are just some of the problems with this approach. For instance, the Oral Physician Program at the University of Kentucky failed because only two students were accepted into the test group, and they lost their sense of belonging to the dental school. We chose the option of comparing the first class or first few classes of the new curriculum with the last class or last few classes of the old curriculum as did several medical schools. However, it must be remembered that many external factors influence the performance of students not participating in simultaneous programs. In our case, a natural disaster, the ice storm that occurred in the province of Quebec in January 1998, forced the cancellation of one week and the reduction of a second week of the dental program at McGill.

Tools for Curriculum Evaluation

A review of the dental literature in English revealed many publications on the evaluation of courses or specific programs, but very limited information on the assessment of an entire curriculum. McCann et al. present the assessment plan for curriculum goals at Baylor College of Dentistry, but do not report results of their various evaluation methods. Although new dental programs are described, their outcomes have not been thoroughly assessed. A closer integration between dentistry and medicine has been advocated by the authors of the IOM report and by others, but to our knowledge, only Login et al. have assessed the effect of such an integration.

The assessment methods selected should link curricular changes to the goals of the program. Indeed, planning curricular changes together with the methods to be used for their assessment is recommended. Unfortunately, changes are often made before assessment of their effectiveness is even considered, which leads to significant limitations in the interpretation of results. A discussion of various proposed curriculum assessment methods follows.

Qualitative Tools

Curriculum Guidelines. In the past, the assessment of the curriculum focused on its process and content. Curriculum guidelines published by various sections of the American Dental Education Association (ADEA) represented a simple way to assess a school curriculum against external standards. These guidelines represented the consensus position of many dental educators about the content of predoctoral dental curricula in specialty areas of dental practice. As a result, the guidelines were conservative and did not reflect the personality of any specific program. Clearly, such guidelines could not be at the leading edge of an educational revolution.

Furthermore, traditional guidelines address curriculum content rather than knowledge transfer and are neither linked to student learning nor linked to competence as a dental practitioner after completion of the program. Nevertheless, linking the content of courses in a curriculum to national guidelines may be a useful exercise and certainly provides some comparative information about the program.

Competency Documents. In the 1990s, the Association of Canadian Faculties of Dentistry, the Commission on Dental Accreditation, and ADEA successfully promoted the change from behavioral and instructional objectives to competency statements. As a result, many dental schools have developed and adopted competency documents or lists of competencies for their students. Competency statements can be the basis for an internal assessment of the curriculum. Ryding has suggested a method for tracking the teaching and assessment of the various competencies through an entire program.

Discussions and Focus Groups. Qualitative assessment can also be accomplished using discussions and focus groups. Discussions usually take place shortly after the completion of courses or major sections of the curriculum and provide the sig-
significant advantage of immediate feedback. Discussions revolve around the courses and cover organization, content, teaching methods, clinical experience, student assessment, and evaluation methods. Participants traditionally include the course director, members of the curriculum committee, student representatives, and the associate dean for academic affairs. Such face-to-face interactions between students and faculty are based on the information gathered through the evaluation questionnaires and have the advantage of clarifying the problems identified and enabling the faculty to respond to criticism. To limit confrontational situations, the evaluations of individual faculty members by students are not part of the discussion. Minutes are taken and later circulated to participants. Discussions are successful assessment methods when both students and faculty feel that their suggestions will be considered and potentially implemented.34

Focus groups, on the other hand, refer to structured discussions of limited duration with a randomly selected group of people.35-37 The moderator of the focus group is generally not a faculty member and has experience in facilitating interactions. The participants are stimulated by the ideas of one another, which is thought to be an advantage. Focus groups can explore thoughts and feelings of participants better than questionnaires. The sessions are audiotaped, and participants remain anonymous. They can be used after completion of the program or during its development for formative assessment to modify the program while in progress.38-39 Focus groups are particularly useful to identify problems and find solutions, but have the major disadvantages of being time-consuming and requiring considerable faculty participation.26 Student-run focus groups provide unique insight regarding student comprehension of material, effectiveness of lecturers, and usefulness of lab time.39

**Teaching Portfolios.** Teaching portfolios are ongoing documents developed by faculty members to gather information on their teaching philosophy, goals and objectives, evaluation techniques, teaching effectiveness, and creativity. An important component is the teacher’s reflection on outcomes. College, nursing, and medical educators have promoted the development of teaching portfolios to evaluate faculty members.40-44 The teaching portfolio is one of the four most frequently used tools to evaluate teachers in North American medical schools.45 Reviewing portfolios of all teachers of a dental school to review an entire program could be extremely time-consuming. Furthermore, portfolios are difficult to assess because their structure and content may vary tremendously.46-48 No dental school, to our knowledge, has reported the use of teaching portfolios for curriculum review. Reece et al.49 feel that portfolios could be utilized to evaluate the quality of teaching and learning in an institution, while Ball et al.50 suggest that the effectiveness of portfolios for curriculum assessment needs to be tested.

**Quantitative Tools**

**Competency Examinations.** Performance-based competency tests are valuable means of demonstrating to the public the standards of care to which the profession is committed.51 They may be the best way to measure the effectiveness of a clinical curriculum, and they have been used as curriculum assessment tools in medical education.52-53 However, some questions need to be answered before they can be adopted as outcome measures. For example: Are the current school’s competency documents comprehensive enough or too detailed? Are all necessary competencies being evaluated? And are the skills being tested a reflection of skills needed in practice? Testing for competency-based education should be done under conditions that simulate problems that dentists face in practice because this provides evidence that the curriculum is achieving its goals.54-55 However, the value of some competency tests is being questioned because the link between these and real clinical skills has never been established.56 For instance, in a survey of 1,265 general dentists, the reported utilization rate for the facebow was 30 percent, 68 percent for custom trays, 59 percent for border moulding, and 51 percent for semi-adjustable articulators.57 Sixty percent of the dentists answered “not essential to the practice of dentistry” as the reason for not using these techniques that are taught in most curricula.

**Board Examinations.** Performance of students on the examinations developed by the National Dental Examination Board of Canada or American National Dental Board may be useful measures of training programs, providing that examination questions fairly cover the content of the curriculum and are consistent with the educational goals of the institution. The primary role of board examinations is to assure society that graduating dentists are competent, which of course is also a major goal of each dental school.26 Matlin22 used performance on the National Dental Board examination as an initial outcome measure of the effectiveness of the new Harvard
dental curriculum although he questions its sensitivity as a tool. The same point was made by Wallace, who feels that a restructuring of these examinations is required to ensure that the entire dental curriculum is properly covered. For instance in Canada, all questions are developed by experts, but the final choice of examination questions is made by general practitioners. This suggests that questions testing new knowledge may have less chance of appearing on the examination than questions relating to old knowledge.

Furthermore, variations in the results between schools can be due to the fact that not all schools give study time prior to the examinations and that some candidates enroll in private preparatory courses. Finally, until very recently, the Canadian Board of Dental Examiners did not release old examination questions, but the students of some schools had built up banks of old examination questions, while student at other schools had not.

The performance of medical students on the Medical Council of Canada Qualifying Examination (MCC), National Board of Medical Examiners (NBME), and United States Medical Licensing Examinations (USMLE) is a frequently reported curriculum assessment tool. However, the value of standardized licensure examinations to determine the success of innovative programs has been questioned for medicine, as it has been for dentistry. It has also been pointed out that the evaluation of the effectiveness of a medical curriculum should not rely solely on standardized tests using multiple-choice questions. Results on board examinations should be reviewed with caution for the following reasons. First, the content of the medical board examination may better reflect a traditional than an unconventional program. Secondly, the multiple-choice format of board examinations is very familiar to students in conventional programs, but varied testing methods are more likely to be used in modified programs. In fact, studies of board examination questions suggest that the format of the questions can influence student performance.

**Oral Comprehensive Examination.** Using an oral comprehensive examination, Login et al. compared the performance of the last class of the traditional lecture-based dental curriculum of the Harvard School of Dental Medicine with the first three classes of the problem-based curriculum, in which the dental students are integrated with the medical students for the first two years. The students from the new curriculum received significantly higher scores in the Science and Medical Knowledge component of the examination and received more positive comments regarding their communication skills. The students were graded using a 5-point scale from honors to failure. The structured oral examination has been reported to result in higher overall and interrater reliabilities than multiple choice questions to assess knowledge acquisition and problem-solving skills of surgical residents.

**Surveys.** Questionnaires are widely used curriculum assessment tools because they provide a lot of information rapidly, at a small cost, and with minimal staff involvement. Questions are generally of the Likert-type with three to seven response categories. Responses take the form of “strongly agree” to “strongly disagree,” “very well prepared” to “very poorly prepared,” “extremely useful” to “extremely useless,” and “unsatisfactory” to “excellent.” Authors report conflicting results on the effect of the direction of the question on responses. Albanese reported that more positive responses are obtained when the positive rating is on the left side, while Barnette found no effect. It is common practice to alternate positively worded stems with negatively worded stems to counteract acquiescent or response set behaviors. Barnette also recommends reversing response alternatives rather than using negatively worded stems.

Visual analogue scales (VAS) have been used in educational research although they are more commonly used to measure symptoms in epidemiological and clinical research. They are continuous scales that have the advantage of being more sensitive to small changes. The simplest VAS is a horizontal line anchored at the end by words like “strongly agree” and “strongly disagree.” The response is a short vertical line drawn somewhere between the two extremes.

**Student Surveys.** Some dental schools conduct surveys of graduating dental students on an annual basis. These surveys provide information on the quality of university and faculty services, practice and postdoctoral education plans, and the adequacy of time allotted to various areas of predoctoral instruction. Educators ask students to rate their level of competence or level of preparation for the program competencies using Likert scales.
provides immediate feedback from graduating students on the entire curriculum.

Assessment of a new curriculum can also take place during a program to evaluate the attainment of short-term learning goals. Authors have compared students’ attitudes to new problem-based curricula with their attitudes to previous traditional courses by asking first- and second-year students of two consecutive years to evaluate the programs.66,68,69,77-80 Wetherell et al.29 reported that the students in the problem-based learning curriculum perceived their workload as more acceptable, giving them more time to think by themselves and with a better balance between theory and practice. These authors reported superior students’ relationships, although Kaufman and Mann12 found that problem-based learning students were less positive about student interactions due to the formation of factions in the class.

Alumni Surveys. Educators feel that students cannot judge the significance of their education until some time after they graduate and have practiced dentistry on their own. Alumni can give significant information on the strengths and weaknesses of the curriculum and on the importance of its various components.66,68,69,77-80 Alumni rate their level of preparation when leaving dental school lower than graduating students, perhaps because they are in a better position to reflect on the level of preparation for practice provided by their education.26,63 Surveys of graduates will need in their future practice. Students’ self-assessment of their level of competency can be compared with conventional curricula.11

Surveys of alumni can also reveal practice patterns, learning behaviors, and levels of satisfaction with the profession, three areas that may provide information about the effectiveness of the curriculum. We must remember that the role of the dental curriculum is not only to develop competence but also confidence and the other attributes of a healthcare professional.81 Mennin et al.64 used alumni questionnaires to show that graduates of a new community-oriented problem-based medical curriculum practiced more frequently in rural and underserved areas. Graduates identified patient problems and curiosity as the motivation behind their learning. It is important to note that attracting graduates to rural areas and providing self-directed lifelong learning skills were two major goals of the new curriculum at the University of New Mexico. Peters29 conducted telephone interviews of Harvard graduates from the New Pathway Program and conventional curriculum. He concluded that behaviors and attitudes in the humanism, life-long learning, and social learning domains can be taught and learned.

However, these surveys have inherent limitations. If the length of time between the end of university training and data gathering is long, it becomes difficult to separate the effects of a curriculum from those of experience.64 The level of interest or lack of interest of alumni for certain aspects of practice can introduce biases in their responses to questions about university programs.65 For these reasons, it has been recommended that surveys be limited to individuals who have graduated in the past ten years.68,69

Evaluation by Instructors. Some authors have used questionnaires filled out by instructors to rate the competency of graduating students.26,76 The performance of students in a hybrid-PBL course in treatment planning with that of students trained in standard lectures was evaluated in this way.73 Directors of postdoctoral general dentistry programs have been asked to rate the level of competency of their own graduates and the overall expected level of competency of general practice residency graduates.62 Surveys have been used to collect instructors’ opinion on student characteristics, methods of instruction, and methods of student evaluation.97 Instructor-based performance assessment would seem to be a useful method to evaluate overall clinical skills, if these are the skills that students will need in their future practice. Students’ self-assessment of their level of competency can be compared with instructor assessments. When this was done in nursing, faculty and students concurred.76

Patient Satisfaction Survey. Traditional clinical teaching in dentistry has been criticized for emphasizing the needs of students rather than those of patients.83 Student-patient interactions are important but difficult to evaluate.81 Patients are able to evaluate communication skills and empathy, so it is logical that a patient-centered curriculum be evaluated by the patients receiving the care. However, there is a tendency for patients to become attached to students, which may lead to artificially high ratings on patient satisfaction questionnaires. To counteract this, standardized patients have been used to rate humanistic skills, history-taking, and examination skills of medical students.54,84-86 Likert-type rating scales are an efficient and reliable way to rate communication skills while checklists are preferable for clinical skills.85 The advantage of using standardized patients is that they can be trained to rate specific skills and can rate the behavior of several students in the same clinical situation. The major disadvantage is that they have to be paid and it takes time and money to train them.26,52
Clinical Productivity. Educators tend to be reluctant to use clinical output or income as an assessment tool. However, there are two justifications for using these variables. First, students need to learn to be efficient, and second, students who are more positive and enthusiastic about their training might be more productive. The advocates of problem-based learning believe that this educational method stimulates the intellectual curiosity of students and is therefore more enjoyable, less stressful, and more productive.21 A comprehensive approach to patient care compared to a requirement-based system has been reported to increase productivity and decrease student stress.87 Increased productivity was taken as a measure of the effectiveness of the new dental curriculum at Harvard while billing data have been used to determine differences in practice profiles between graduates of problem-based and conventional medical programs.22,53

Conclusion

Dental schools need to continuously review and modify their programs to ensure that they reflect the changing oral health needs of society as well as developments in knowledge and in clinical practices. The effectiveness of a curriculum cannot be measured using a single evaluation method because no one tool can capture all the information required. Several methods and groups of evaluators are needed to provide a more global perspective on the entire curriculum. To study the effectiveness of the new curriculum on academic and clinical performance at McGill, we decided to use five evaluation tools: clinical productivity, instructor surveys, graduating student surveys, patient surveys, and board examinations. These five tools were selected because they required reasonable faculty and financial resources.

This evaluation had the advantage of combining external assessment tools (board examinations and patients’ surveys) with internal tools (clinical productivity, instructors’ surveys, and student surveys). We felt that it was important to ask for the opinions of the students who receive the education, the patients who receive the care, and the instructors who monitor the progress of the students. Another advantage was that the five tools adopted were quantitative tools and were therefore less subject to interpretation. The results of the assessment of the McGill curriculum will be discussed in “Assessing the Effectiveness of a New Curriculum: Part II,” which will be published in an upcoming issue of this journal.

REFERENCES


80. Woodward CA, Ferrier BM. The content of the medical curriculum at McMaster University: graduates’ evaluation of their perception for postgraduate training. Med Educ 1982;17:54-60.


