Assessment of Education and Research in Biomedical Informatics

J.H. van Bemmel
Erasmus University and Erasmus Medical Centre, Rotterdam, The Netherlands

Summary

Objectives: The existence and survival of university institutes is increasingly dependent on assessments of research and education. In many countries also departments of biomedical informatics are assessed at regular intervals, often as part of the review of a Medical or Health Sciences Faculty, or a Research School. The article underlines the importance of periodic evaluation of research and education in biomedical informatics.
Methods: Quality assessment, if done by an independent review committee of peers, is a suitable instrument to obtain insight into the quality and accountability of both education and research. Key instruments for the assessment of education and research are well-defined protocols that are used for self-assessment. These self-assessment reports form the inputs for the independent review committee.
Results: The outcomes of the assessments are directly related to the quality of research, which is visible in publications in peer-reviewed journals. Internal quality management tools contribute to a large extent to the improvement of the quality of education and research.
Conclusion: External assessment – review by peers – is increasingly used as the final step of an integral quality system for research and education. This is particularly important if the results of biomedical informatics R&D are to be applied in clinical practice. A positive outcome of an assessment can only be expected from a long-term investment in the quality of research and researchers who publish their results in peer-reviewed journals.
Haux R, Kulikowski C, editors. IMIA Yearbook of Medical Informatics 2006. Methods Inf Med 2006; 45 Suppl 1: S5-10.

Keywords

Research, education, biomedicine, assessment, peer review, quality management

1. Introduction

Biomedical informatics[1] is both an applied and a more or less basic discipline. Research and development (R&D) in biomedical informatics is done by industry, but for the most part the advancement of the field takes place in universities and institutes of higher education. For the quality of industrial R&D, guidelines and protocols have been established, such as the ISO 9000 norms [1]. R&D and education at universities are most of the time assessed by review committees, consisting of independent researchers, such as peers [2]. Increasingly, the existence and survival of research institutes, including those of biomedical informatics, depends on assessments that are done at regular intervals of, say, 4 to 5 years. In many countries, the assessment of university departments of biomedical informatics is done in the framework of the review of an entire Medical Faculty or a Research School. This is also the situation in the Netherlands, where the author has frequently been involved in the assessment of biomedical research and education in the widest sense. Therefore, this introductory article is largely based on his own experience, both in conducting research in his own Institute and his participation in assessments of university research in several European countries.

The intention of this article is to underline the importance of the evaluation of research and education in biomedical informatics. The outcome of these evaluations is directly related to the quality of research projects. A positive judgment can only be expected if the researchers involved have made a long-term investment in the quality of their research, made visible by publications in peer-reviewed journals.

There is another aspect worth mentioning when assessing one's research. It pertains to the appropriate evaluation of the methods and systems that are used and developed in our research. This is particularly important if the results of biomedical informatics R&D are to be applied in clinical practice. Although this aspect of evaluating one's own R&D is not the prime goal of this article, the outcome of such evaluations – e.g., done by conducting clinical trials [3] – will positively contribute to the outcome of the other type of assessment described above. Technology assessment in health care falls into this category as well [4]. Some examples of the latter type of evaluation studies, conducted in biomedical informatics worldwide, can be found in the publications of many research institutes, e.g., at Stanford University [5], Columbia University in New York [6], Vanderbilt University in Nashville [7], the University

[1] In this article we use the term 'biomedical informatics' to comprise related terms, such as medical or health informatics, medical or health information processing, nursing informatics and other similar fields.

of Columbia-Missouri [8], Regenstrief Institute/Indiana University [9, 10], McMaster University in Hamilton [11, 12], Erasmus University Rotterdam [13, 14, 15], or the Universities of Heidelberg, Amsterdam or Innsbruck [16] – but the list is much longer. Ideally, such evaluation studies are carried out by an independent party, and are at least conducted as much as possible in an objective way.

2. Goals and Scope of the Assessment

The goals of assessing the quality of education and research are twofold:
1. Improvement of education and research;
2. Accountability of how resources were spent.

We should realize that the rating of the quality of education and research is a relative matter, even when formal procedures are applied. There are no absolute standards. Moreover, we should be aware that assessment is done by human beings – hopefully knowledgeable peers – who are undoubtedly influenced and limited by their own background and experience. Only by mutual comparison with similar institutions elsewhere – hopefully with the best – can we obtain an impression that is as objective as possible of the position of an institution on the quality scale.
In most countries, education and scientific research are evaluated separately. Besides, each of the approximately 40-50 different disciplines[2] in the biomedical field is almost always assessed independently. After about 25 years of experience with assessment of education and research in our country, scientific research is no longer evaluated solely on the basis of scientific mono-disciplines, but predominantly along the lines of major research programs and research schools. The reason for this is that, nowadays, in many areas, scientific research has a multidisciplinary character [17].
In the following we will describe the main elements of the evaluation of education and research in universities, as said, primarily based on the own experience of the author.

3. Assessment of University Education

Striving for high-quality university education is in the first place advantageous for the students themselves, but also of great value for society, including the future employers of our students. In the Netherlands, reports on the quality of education are intended for the university, future students, and the Government, which provides part of the financial resources and oversees the quality of higher education. Assessment of higher education is conducted in many countries (see, for instance [18]). The way in which the position of an institution is located on a quality scale is important. Different ways of assessing this quality include:
1. asking the opinions of the current student generation;
2. interviewing postgraduates;
3. asking the opinions of colleagues at other universities;
4. requesting a self-assessment from the faculty staff;
5. performing an external peer review.

All such types of evaluations are being carried out in our country:
1. The opinions of the students are requested on a regular basis, usually by the university itself. In addition, a countrywide magazine annually collects the opinions on university education by interviewing students;
2. The opinions of postgraduates are collected by means of a national labor market monitor;
3. The same magazine as in 1. also requests professors at other universities to give their opinions on education offered at other universities;
4. Formal self-assessment studies according to a prescribed protocol are obligatory every 6 years and serve as input for an external quality assessment;
5. Peer reviews, based on these self-assessments, are organized by an independent quality assurance agency.

Not surprisingly, the outcomes of all such evaluations reveal different results. However, the different viewpoints by students and peers do not invalidate the formal assessments. On the contrary, they may be used to throw a different light on the matter. The self-assessment reports form the basis for site visits and peer reviews, in which the quality of comparable courses offered by different universities are compared.

Peer Review

All evaluations contain subjective elements, and even with judgments from peers one should be cautious, because the composition of a review committee may influence the outcome. This is why members of a review committee should be drawn from different centers. Moreover, it is advisable to use a well-defined protocol with clear instructions pertaining to the intention of the assessment, to exclude individual opinions and to enable comparisons over time. In short, the task of an independent review committee is to judge the quality, to give advice, to compare with other institutions, and to report to the university and the government.

[2] Think of separate disciplines such as anatomy, biochemistry, genetics, internal medicine, neuroscience, oncology, public health, radiology, and biomedical informatics. The number of disciplines is, of course, dependent on a coarse or fine subdivision.

Improvement

The self-assessment is intended to report on the strengths and the weaknesses of the educational process. The external peer review is in fact a meta-analysis of the self-assessment reports, using the expertise of the referees. If possible, the review committee pays site visits to the institutes participating in the assessment, it offers recommendations, and explains which educational aspects should be improved. Such recommendations are reviewed at a future assessment in order to judge whether any actions were taken and improvements achieved. It should be realized that a review committee report is typically published about one year or more after the self-assessment report was composed, so that improvements may already have been set in motion and possible negative judgments may already be outdated.

Accountability

Accountability is an important aspect of all assessments. Because, traditionally, in many countries universities are at least partly publicly funded, they are required to report to the government on the quality of the education they offer and on how they have spent their resources. Implicitly, such reports are also intended to offer information to future students and their parents and society at large, taxpayers included.

Assessment rounds

In our country, universities have completed several rounds of the external assessment of university education, rounds with intervals of 5-6 years, a period in which all courses are reviewed. Renewal of the most recent assessment protocol was set in motion because of the continuously changing demands from society, shifts from public to private funding, the introduction of information and communication technology in the curricula, and the introduction of the bachelor-master structure in the European Union. The latter intends to bring all higher education to an internationally comparable level.

4. Assessment of Scientific Research

Assessment of scientific research at our universities was started in the early 1980s. External review committees, mainly consisting of non-Dutch members, carry out the evaluation on the basis of an assessment protocol. In the past, the assessments for biomedical and health sciences research were carried out under the auspices of the Royal Netherlands Academy of Arts and Sciences (KNAW) [19]. In 2001, a renewed assessment strategy was defined, partly based on the model developed by the European Federation for Quality Management [20]. This new assessment system is no longer directed towards research carried out within separate disciplines, but concerns large, comprehensive research clusters, such as the neurosciences or the health sciences. Biomedical informatics is part of the latter cluster. Self-assessment reports are to be written every three years, and are reviewed by an external international review committee every six years.

Accountability

Scientific research at Dutch universities is financed by different sources:

- Public funding by the Government (the Ministry of Research and Education), as part of the budgets that universities obtain for education and research;
- Highly competitive funding from the Netherlands Organization for Scientific Research [21] (NWO);
- Subsidies from other research funds (e.g., for cardiac diseases or cancer, generally charities);
- Competitive funding by the European Union (e.g., via its Framework Program 6 [22]);
- Other Ministries (e.g., for Health or Economic Affairs);
- Industrial companies and other private corporations.

The balance between public research funding (the first source above) and all other funding resources is shifting towards the latter, with percentages of 50/50 or even 30/70 not being unusual. Of course, all funding organizations are very interested in the quality of research carried out in universities. However, it should be realized that universities and their scientific staff are also increasingly involved in semi-commercial enterprises, funded privately. This research and their outcomes are often not accessible to open assessment, because of the commercial interests of the parties involved. The following description, therefore, pertains mainly to the assessment of research, financed by direct and indirect public funding and open for assessment.

Self-assessment Reports

Self-assessment reports at the level of the Institute to be assessed should contain several elements, such as:

1. *Characterization of the institute*. Mission and research; collaborations and affiliations.
2. *Leadership*. Organizational structure; names of directors, departmental heads; list of research programs and program leaders.

3. *Research strategy*. Organizational context; plans for the short term and the long term.
4. *Researchers*. Personnel policy: recruitment, selection, training, career planning, mobility; list of research staff, function, tenure, and period of appointment.
5. *Resources and funding*. Financial situation: funding and expenditures; research contracts; research infrastructure and capital investments; future funding prospects.
6. *Processes to support research*. Research culture, such as teamwork, communication and exchanges with other institutes, supervision of PhDs, internal quality assurance.
7. *Reputation*. Academic reputation expressed in, e.g., citation scores, prizes and awards.
8. *Internal assessment*. Monitoring of research management.
9. *Research outcomes*. Publications in different categories: in refereed journals, as scientific papers, in books, as monographs, and as professional products; patents.
10. *External appreciation*. Effect of the dissemination of research outcomes.
11. *Future perspectives*. See the description below, under Research management.

For all research programs, such data have to be provided and, in addition, five key publications best characterizing the research, three in full text are to be included. All other publications are listed according to the categories mentioned under point 9 above.

External peer review

In the external assessment, four aspects are to be considered:

- Quality of the scientific research;
- Productivity of the scientific output;
- Relevance of the research for academia and society;
- Future perspectives, feasibility and vitality of the research.

Although the report of the peer review committee should contain judgments on research programs in wording, it has also been proven useful to position the research on a 5-point qualitative scoring scale, which are given for all four aspects separately:

5. *Excellent*. Research that is internationally at the forefront and has a high impact.
4. *Very good*. Research that is internationally competitive and nationally at the forefront.
3. *Good*. Research that is nationally competitive and internationally visible.
2. *Satisfactory*. Research that is solid but not exciting. Nationally visible.
1. *Unsatisfactory*. Research that is neither solid, nor exciting. Not worth pursuing.

In Table 1 we give an example of the outcome of an assessment process. We present the outcomes for all biomedical research in our country, covering the 5-year period before 1998. This outcome was the final judgment of an international review committee of peers, chaired by the author of this article. The Table summarizes the scores that were given to Medical Faculties (Nrs. 1-9) and large biomedical research Institutes (Nrs. 10-18). In this overview all research projects within a Faculty or Research Institute were clustered into one of 14 main research themes, such as Genetics, Immunology, Neurosciences, Cardiovascular Diseases, Oncology, and Health Sciences. Biomedical Informatics belonged to the last cluster. The Table shows how the Faculties and Institutes scored for the main research themes. For a main research theme that was given an overall quality score of, e.g., 3-4, 0.5 points were given to score 3 and 0.5 points to score 4. Although no scores of ‘unsatisfactory’ were given, a judgment with scores 2 and even 3 might have led to consequences for the research projects and the researchers concerned. It can be seen that 67.5 of a total of 114 (60%) was judged as very good or excellent.

Research schools

Research schools are regularly assessed as well. It was decided that KNAW should be designated as the coordinating organization for the accreditation of all 110 research schools. An accreditation period lasts 5 years. The research schools, covering the entire field of research at universities in our country, have all been reviewed and accredited, some for the third time. Many departments, earlier operating along lines of individual disciplines, are now part of research schools, which often bear an interdisciplinary character. This also applies to the field of biomedical informatics.

Research management

In addition to the evaluation of scientific research, the review committee is requested to assess research management, also based on self-assessment reports. Aspects of research management that are reviewed are the following:

- *Management structure*. How the internal management structure for research is shaped and how and by whom decisions on research are made is assessed. The collaboration between a research institute and the university hospital is also reviewed. Nowadays, in all our universities the Medical Faculty has been integrated with the University Hospital, governed by one Board.

Table 1 Outcome of the assessment of all biomedical research in the Netherlands in the year 1998, covering the 5-year period before. The Table lists the number of research themes in a Faculty or Institute with scores from 1 (low quality) to 5 (high), based on the final judgment of an international review committee of peers. For further explanation, see the text.

Medical Faculty or Research Institute	SCORE 5	4	3	2	1
1	-	3	2	-	-
2	4	2.5	4	0.5	-
3	-	5.5	3.5	-	-
4	-	6	5	-	-
5	3.5	4	2.5	-	-
6	1.5	3.5	4	1	-
7	1	5	-	-	-
8	0.5	3.5	5	-	-
9	1.5	2.5	2	1	-
10	-	1.5	0.5	1	-
11	1.5	2.5	-	-	-
12	1.5	0.5	1	-	-
13	-	-	1	-	-
14	-	-	0.5	0.5	-
15	-	3	1.5	0.5	-
16	-	1	3	-	-
17	4	1	-	-	-
18	-	3.5	7.5	-	-
Total	19	48.5	42	4.5	-

- *Financial management*. The report should contain funding resources available and how budget allocations take place. It is important to review whether funding is available for maintaining continuity of research, e.g., by the presence of a revolving fund.
- *Infrastructure and support*. This concerns the available infrastructure for research, the regular renewal of equipment and its maintenance status. Also the support by information technology, both for research and the library, are to be assessed.
- *Human resources*. This aspect concerns the availability of human resources and the composition and turnover of staff, and how and by whom new staff is appointed. The planning of careers and the training of personnel is essential. Age and gender aspects are also to be addressed.
- *Internal assessment of the quality of research*. The existence of a system for internal quality assurance of research is important for management decisions. Assessment is made of whether negative outcomes have an effect on the discontinuation of research, but also whether excellent research is being stimulated and rewarded.
- *Ethical issues*. Because of the specific character of biomedical research, ethical aspects have also to be addressed. For instance, whether the appropriate committees were installed for both human and animal-related research is included, as are issues pertaining to privacy and data protection.

5. Conclusions

As we have discussed, the regular assessment of institutes of higher education and research, including those of biomedical informatics, is of pivotal importance for the continuation of their existence. In many countries departments of biomedical informatics are assessed at regular intervals, often as part of the review of a Medical or Health Sciences Faculty, or a Research School. The outcome of these assessments is strongly dependent on the quality of research, which is reflected in publications in peer-reviewed journals.

Based on the preceding overview, we now list the main conclusions on assessing education and scientific research at universities:

1. Quality assessment, if done at regular intervals and by an independent review committee of peers, is a suitable instrument to obtain insight into the quality and accountability of both education and research.
2. Internal quality management tools contribute to a large extent to the improvement of the quality of education and research.
3. Key instruments for the assessment of education and research are well-defined protocols that are used for self-assessment.
4. External assessment – review by peers – is increasingly used as the final step of an integral quality system for research and education.
5. Because scientific research in many domains, including that of biomedical informatics, bears a multi-disciplinary character, the assessment itself should also be carried out by a review committee consisting of members from different disciplines.
6. A positive outcome of an assessment can only be expected from a long-term investment in the quality of research and researchers who publish their results in peer-reviewed journals.

References

1. See http://www.iso.org/iso
2. See, e.g., http://www.hero.ac.uk/uk/research; http://www.qanu.nl; http://www.stratresearch.se; http://www.inserm.fr/fr/inserm/evaluation; http://www.dfg.de/ranking
3. See http://www.fda.gov/oc/gcp
4. See http://www.htai.org
5. Yu VL, Buchanan BG, Shortliffe EH, Wraith SM, Davis R, Scott AC, et al. Evaluating the performance of a computer-based consultant. Comp Progr Biomed 1979; 9: 95-102. (This is early work, done in Stanford)
6. Shea S, DuMouchel W, Bahamonde L. A meta-analysis of 16 randomized controlled trials to evaluate computer- based clinical reminder systems for preventive care in the ambulatory setting. J Am Med Inform Assoc 1996; 3: 399-409.
7. Miller RA, Pople HE Jr, Myers JD. INTERNIST-1, An Experimental Computer-based Diagnostic Consultant for General Internal Medicine. N Engl J Med 1982; 307: 468-76. (This is early work, done in Pittsburgh)
8. Krishna S, Balas EA, Spencer DC, Griffin JZ, Boren SA. Clinical trials of interactive computerized patient education: implications for family practice. J Fam Pract 1997; 45(1): 25-33.
9. Tierney WM, McDonald CJ. Testing informatics innovations: the value of negative trials. J Am Med Inform Assoc 1996 Sep-Oct; 3(5): 358-9.
10. Dexter PD, Wolinsky FD, Gramelspacher GP, Zhou XH, Eckert GJ, Waisburd M, et al. Effectiveness of computer-generated reminders for increasing discussions about advance directives and completion of advance directive forms. A randomized, controlled trial. Ann Intern Med 1998; 128: 102-10.
11. Hunt DL, Haynes RB, Hanna SE, Smith K. Effects of computer-based clinical decision support systems on physician performance and patient outcomes: a systematic review. JAMA 1998; 280: 1360-1.
12. Garg AX, Adhikari NK, McDonald H, Rosas-Arellano MP, Devereaux PJ, Beyene J, et al. Effects of Computerized Clinical Decision Support Systems on Practitioner Performance and Patient Outcomes. A Systematic Review. JAMA 2005; 293: 1223-38.
13. Willems JL, Abreu-Lima C, Arnaud P, van Bemmel JH, Brohet C, Degani R, et al. The diagnostic performance of computer programs for the interpretation of electrocardiograms. N Engl J Med 1991; 325: 1767-73.
14. Van der Lei J, Van der Does E, Man in 't Veld AJ, Musen MA, Van Bemmel JH. Response of general practitioners to computer-generated critiques of hypertension therapy. Methods Inf Med 1993; 32: 146-53.
15. Van Wijk MA, van der Lei J, Mosseveld M, Bohnen AM, van Bemmel JH. Assessment of decision support for blood test ordering in primary care. A randomized trial. Ann Intern Med 2001; 134: 274-81.
16. Ammenwerth E, De Keizer N. An inventory of evaluation studies of information technology in health care: Trends in evaluation research 1982-2002. Methods Inf Med 2005; 44: 44-56.
17. See, e.g., National Academy of Science. *Facilitating Interdisciplinary Research*. Washington DC: National Academy Press, 2004.
18. http://www.ncrel.org/sdrs/areas/stw_esys/4assess.htm
19. See http://www.knaw.nl
20. See http://www.efqm.org
21. See http://www.nwo.nl
22. http://europa.eu.int/comm/research/fp6

Correspondence to:
Jan H. van Bemmel, PhD
Professor of Medical Informatics
Erasmus Medical Center and Erasmus University Rotterdam
P.O. Box 1738
3000 DR Rotterdam
The Netherlands
E-mail: J.vanBemmel@ErasmusMC.nl