Evaluating Quality in Clinical Cancer Research: The M.D. Anderson Cancer Center Experience

James D. Cox^a Sergio A. Giralto^b Mary L. Veazie^c Jaffer A. Ajani^d Janet M. Bruner^e
Ka Wah Chan^f Walter N. Hittelman^g Kelly K. Hunt^h Revathy B. Iyerⁱ Daniel D. Karp^j
Deborah A. Kuban^a Scott M. Lippman^k Issam I. Raad^l Miguel A. Rodriguez-Bigas^h
Leonard A. Zwelling^{g,m} Maurie Markmanⁿ

^aDivision of Radiation Oncology, ^bDepartment of Blood and Marrow Transplantation, ^cAcademic Programs, ^dDepartment of Gastrointestinal Medical Oncology, ^eDepartment of Pathology, ^fDepartment of Pediatrics, ^gDepartment of Experimental Therapeutics, ^hDepartment of Surgical Oncology, ⁱDepartment of Diagnostic Radiology, ^jDepartment of Thoracic/Head and Neck Medical Oncology, ^kClinical Cancer Prevention, ^lDepartment of Infectious Diseases, ^mResearch Administration, ⁿDepartment of Gynecologic Medical Oncology and Clinical Research, The University of Texas M.D. Anderson Cancer Center, Houston, Tex., USA

Key Words

Cancer centers · Academic medical centers · Clinical trials

Abstract

Background: Despite the unquestionable importance of clinically oriented research designed to test the safety and efficacy of new therapies in patients with malignant disease, there is limited information regarding strategies to evaluate the quality of such efforts at academic institutions. **Methods:** To address this issue, a committee of senior faculty at the University of Texas M.D. Anderson Cancer Center established specific criteria by which investigators from all departments engaged in clinical research could be formally evaluated. Scoring criteria were established and revised based on the results of a pilot study. Beginning in January 2004, the committee evaluated all faculty involved in clinical research within 35 departments. Scores for individual faculty members were assigned on a scale of 1 (outstanding) to 5; a score

of 3 was set as the standard for the institution. Each department also received a score. The results of the evaluation were shared with departmental chairs and the Chief Academic Officer. **Results:** 392 faculty were evaluated. The median score was 3. Full professors more frequently received a score of 1, but all faculty ranks received scores of 4 and 5. As a group, tenure/tenure track faculty achieved superior scores compared to nontenure track faculty. **Conclusions:** Based on our experience, we believe it is possible to conduct a rigorous consensus-based evaluation of the quality of clinical cancer research being conducted at an academic medical center. It is reasonable to suggest such evaluations can be used as a management tool and may lead to higher-quality clinical research.

Copyright © 2009 S. Karger AG, Basel

The relevance of clinical investigation as an essential component in the advancement of the health of a population is well established. Much has been written regarding the relative value of different levels of evidence, from the major importance of the prospective phase III randomized trial and meta-analysis of multiple trials to the limited utility of individual case reports or retrospective institutional reviews [1, 2].

However, such efforts have focused on the final products of clinical research rather than on the process used to conduct the research. Unfortunately, recent publications have documented the extent of failure to publish research that had been presented at national meetings [3], including results from phase III prospective clinical trials in cancer [4]. Such publications suggest the need to evaluate the entire process of clinical research, with the goal of improving the quality of that process, ultimately for the benefit of patients by finding safe and effective new treatments and eliminating ineffective therapies.

Many academic medical centers have large clinical research enterprises. Although patient care revenues may have funded clinical research in the past, current reimbursement schedules do not accept research as a component of standard patient care. Thus, not surprisingly, department chairs and senior research administrators in academic health centers who were surveyed about the quality and health of clinical research cited increased pressure to see patients, insufficient clinical revenues and lack of external support for clinical research as serious issues [5]. Adding to this concern, it is appreciated that national funding agencies place a priority on laboratory research and provide relatively little support for clinical investigations except for those that are carried out by multi-institutional cooperative groups. As a result, a large percentage of clinical investigation in cancer is supported by industry, primarily large pharmaceutical companies [6].

The University of Texas M.D. Anderson Cancer Center (MDACC), like most academic medical centers, performs a large number and a wide variety of clinical investigations. These range from translational laboratory-based initiatives to phase I through phase III clinical trials. Financial support for these efforts is provided by peer-reviewed funding (e.g., National Cancer Institute, Department of Defense), industry, and MDACC itself. The majority of faculty involved in the care of patients at MDACC participates at some level in clinical research.

In 2004, more than 30,000 patients participated in MDACC Institutional Review Board (IRB)-approved protocols (interventional and laboratory-based studies).

Major resources of the institution are committed to the clinical research enterprise, and it is appropriate that the institutional leadership should be concerned with the effective and efficient use of these resources.

The hypothesis was advanced that the impact and influence of clinical research by a faculty whose patient care is avowedly 'research driven' could be evaluated. We describe here the development of this process and the initial results of this novel effort.

Methods

At the request of the President and the Chief Academic Officer (CAO) of MDACC, a committee of faculty members (Clinical Research Impact Committee, CRIC) was created to develop a means of evaluating clinical research conducted at the institution. The CRIC is a subcommittee of the Research Council (comprised of basic science and clinical leaders) and reports to that Council and to the CAO. Committee members are selected by the Chair of the CRIC and the CAO from senior faculty members with extensive clinical research experience. In its initial make-up, the CRIC consisted of 12 professors and 3 associate professors representing all disciplines within the institution involved in the clinical research enterprise (e.g., surgery, radiation oncology, medical oncology, pediatric oncology, internal medicine subspecialties, pathology, radiology, cancer prevention, translational laboratory-based).

The committee defined 'clinical research' as activities requiring approval by the MDACC IRB. The committee developed a form to be completed by each faculty member, the key elements of which included: (a) self-assessment by the faculty members of their most important contributions to clinical research (limited to 100 words); (b) peer-reviewed research funding; (c) research support from all other sources (external and internal peer-reviewed sources, industry, philanthropy); (d) publications [total over 5 years, the 3 most important papers (in their personal opinions), and the 3 most frequently cited publications (using the ISI Global Net Website) over the previous 7 years]; (e) collaborative activities (interdepartmental, multi-institutional, national and international), with an emphasis on leadership of those collaborations and on publications resulting from them; (f) protocol leadership and publication of results from protocols, and (g) archival and database-related projects. The individual faculty member's NIH biosketch was attached to provide additional information and clarity.

The CRIC also established scoring criteria (table 1), with the intent that scores be assigned to each faculty member on the basis of their self-assessments and data from the M.D. Anderson Office of Protocol Research. This office maintains a computerized database of all protocols, past and present, both prospective and archival (including those derived from institutional databases).

Scores ranged from 1 (outstanding) to 5. The convention adopted, and rigorously adhered to by the CRIC, was that a score of 3 would be considered the standard within the institution to permit the possibility of superior performance by faculty to be scored as 2 or 1 and inadequate performance to be scored as 4 or 5. The global score represented the consensus of the committee and was usually but not necessarily a numerical average of the scores for

Table 1. Clinical research impact committee guidelines for assessing individual quality score

Score	
Research impact and relevance	
1, 2	Research has established new directions in a field or has moved a field forward
3	Research is making contributions to an important field of investigation
4, 5	Research has little recognizable impact
Research funding	
1, 2	Excellence in peer-reviewed funding or sustained other extramural funding for multiple projects as principal investigator
3	Extramural funding as principal investigator
4, 5	Insignificant peer-reviewed or other funding as principal investigator
Publications and patents	
1, 2	Continuous record of highly cited publications in high-impact journals
3	Cited publications in recognized specialty journals
4, 5	Infrequently cited publications in low-impact specialty journals
Collaborative activities	
1, 2	Leadership role in one or more institutional or national/international collaborative projects resulting in highly cited publications in high-impact journals
3	Active participation in one or more collaborative projects
4, 5	No significant participation in collaborative projects
Protocol status and related abstracts	
1	Chair of at least two completed protocols, with publications, that have each accrued at least 20 patients
2	Chair of at least one completed protocol, with publications, that has accrued at least 20 patients
3	Co-chair of at least two completed protocols, with abstracts
4	Co-chair of at least one completed protocol, without abstracts or publications
5	Neither a chair or co-chair of a protocol, with no protocols or publications
Retrospective, archival or other databases	
1	Chair of two studies of 100 or more patients
2	Chair of one study of 100 or more patients
3	Chair of one study of 50–100 patients
4	Chair of one study of less than 50 patients
5	No study as chair

The following guidelines for individual quality scores in all categories represent sliding scales that increase or decrease with respect to rank and time in rank.

Fig. 1. Distribution of global scores for faculty members. The distribution of global scores among 392 clinical research faculty members, as assessed by the Clinical Research Impact Committee. Note that 30 faculty members were scored with a ‘P’ for ‘potential’ indicating that their length of state as an M.D. Anderson faculty member was <3 years.

the individual criteria listed in table 1. Finally, the committee established a score for each department and presented that score, with suggestions for improvements (e.g. developing a formal research plan, mentoring of junior faculty) to the CAO and Vice President for Clinical Research who in turn met with each department chair to discuss the results.

The CRIC met weekly. Consensus was reached on scores for each component of the evaluation instrument for each faculty member, department by department. Faculty were excluded from evaluation if they had been at MDACC for less than 2 years. Faculty members who had been with the institution from 2 to 3 years could be given a score of ‘P’ for potential, but no other score was given. Physicians whose efforts were devoted entirely to the care of patients (where they were not listed as a principal investigator on any IRB-approved protocol) or those who exclusively pursued laboratory investigations were also excluded from this evaluation.

The CAO and the Vice-President for Clinical Research reviewed the results and the committee’s recommendations with the department chairs. Department chairs could review individual scores with faculty members. The scores were not intended to be used for promotion or tenure decisions, but rather as a management tool by the chairs and institutional leadership on matters such as resource distribution, space allocation, and new recruitment opportunities.

Results

At the time of this evaluation, over 500 faculty members of MDACC were involved in the care of patients of whom 392 met the criteria noted above (i.e., faculty ap-

Fig. 2. Distribution of global scores by faculty rank.

pointment for 3 or more years and involved in clinical research). The distribution of global scores for these faculty members is shown in figure 1. As expected from the scoring method, the most common score was 3. Faculty with global scores of 1 or 2 were considered to have outstanding or superior performance, and global scores of 4 or 5 reflected performance considered below the standard for the institution. A score of 'P' was assigned to 30 faculty members (7.6% of total evaluated, all of whom had been at MDACC for less than 3 years), indicating that they may make substantial contributions in the future, but that assigning a score would be premature at the time the evaluation was completed. The distribution of global scores by faculty rank is shown in figure 2. Scores of 1 or 2 were found more often among more senior faculty although scores for some senior faculty were 4 or 5. Most faculty with 'P' scores were assistant professors. Tenured and tenure track faculty (fig. 3) were ranked more favor-

ably than nontenure track faculty (fig. 4). Department rankings are shown in figure 5.

Discussion

The ultimate goal of clinical research is to improve the practice of medicine and clinical outcomes, and to lay the foundation for the conduct of future health-related investigations. The types of evidence that have affected cancer patient care throughout the past century have been reviewed [7]. Although the quality of the evidence resulting from clinical research has been evaluated and ranked [1], we were unable to find published reports of attempts to evaluate the research process itself with a view toward improving it.

One of the problems with the clinical research process that has garnered attention recently has been the disqui-

Fig. 3. Distribution of global scores for tenured/tenure track faculty by rank.

eting absence of a direct link between the presentation of clinical research results in abstract form at national scientific meetings and the subsequent appearance of full reports in the peer-reviewed literature. Krzyzanowska et al. [4] compared the number of publications that appeared after oral presentations given at annual meetings of the American Society of Clinical Oncology between 1989 and 1998. They found that 26% of oral presentations describing large ($n \geq 200$) phase III comparative clinical trials had not been published 5 years or more after presentation. More of the studies with positive results were published (81%) than were those with negative or neutral results (68%), but 19% of trials with statistically significant results remained unpublished 5 years after their presentation. In a similar review of publications after oral presentations given at the 1995 annual meeting of the Radiological Society of North America, only one third of the presentations had been followed by published manu-

scripts between 1996 and 2000 [3]. A meta-analysis of publication records of studies that previously had appeared as abstracts found a mean publication rate of 45% among 46 reports [8].

Several ethical issues arise from prospective trials not being published, among them the promise to, and expectation of, research subjects who consent to participate in trials that the results of the study will be analyzed and published in the peer-reviewed literature, hopefully for the benefit of others [9]. It is noteworthy and unfortunate that publication of results of prospective clinical investigations has not been considered a major requirement in discussions of clinical research ethics [10].

The results of our study were sufficient to demonstrate to both the CRIC membership and the MDACC leadership that evaluation of clinical research was both possible and valuable. In fact, at the time of the writing of this report, the institution is undertaking a second review of

Fig. 4. Distribution of global scores for nontenure track faculty by rank.

Fig. 5. Distribution of overall scores for entire departments engaged in clinical research.

the clinical research enterprise (initiated approximately 3 years after the initial review). Whether the results will lead to improvements in the scores of the faculty involved in clinical research will not be known until a second round of evaluations is completed, so that each faculty member and each department can be compared with the baseline results from the first round.

However, the benefits already derived from this effort have been an increased focus of an ongoing discussion regarding the quality of clinical research, establishment of agreed-upon criteria for excellence in this endeavor, and determining baselines for future benchmarking of both individuals and departments.

It is relevant to acknowledge that there were some concerns raised by the faculty regarding the conduct of this review. However, full explanations of the process and identification of the CRIC committee members led to general acceptance of, and need for, this institutional ef-

fort. A critical component of gaining this acceptance was clearly segregating this activity from any other evaluative process (e.g. on promotion or tenure).

Finally, it must be acknowledged that the process of evaluating clinical research quality itself could be improved. This is but a first attempt and may eventually be considered rudimentary. If these data promote discussions that eventually lead to better clinical research that improves the care of patients in the nation and around the world, it would be considered a fruitful beginning.

Acknowledgments

Supported in part by Grant CA 16672 from the National Cancer Institute, National Institutes of Health, US Department of Health and Human Services.

The authors wish to acknowledge the support and encouragement provided by Margaret Kripke, PhD, Executive Vice President and Chief Academic Officer, University of Texas M.D. Anderson Cancer Center, without which this entire effort would not have been possible.

References

- 1 Swedish Council on Technology Assessment in Health Care: Radiotherapy for cancer: a critical review of the literature. *Acta Oncol* 1996;35(suppl 7):1-152.
- 2 Nygren P, Glimelius B; SBU-Group: The Swedish Council on Technology Assessment in Health Care (SBU) report on cancer chemotherapy-project objectives, the working process, key definitions and general aspects on cancer trial methodology and interpretation. *Acta Oncol* 2001;40:155-165.
- 3 Arrive L, Boelle PY, Dono P, et al: Subsequent publication of orally presented original studies within 5 years after 1995 RSNA Scientific Assembly. *Radiology* 2004;232:102-106.
- 4 Krzyzanowska MK, Pintilie M, Tannock IF: Factors associated with failure to publish large randomized trials presented at an oncology meeting. *JAMA* 2003;290:495-501.
- 5 Campbell EG, Weissman JS, Moy E, et al: Status of clinical research in academic health centers. *JAMA* 2001;286:800-806.
- 6 Bodenheimer T: Uneasy alliance. Clinical investigators and the pharmaceutical industry. *N Engl J Med* 2000;342:1539-1544.
- 7 Cox JD: Evidence in oncology: the Janeway Lecture 2000. *Cancer J* 2000;6:351-357.
- 8 Scherer RW, Dickersin K, Langenberg P: Full publication of results initially presented in abstracts: a meta-analysis. *JAMA* 1994;272:158-162.
- 9 Pich J, Carne X, Arnaiz JA, et al: Role of a research ethics committee in follow-up and publication of results. *Lancet* 2003;361:1015-1016.
- 10 Emanuel EJ, Wendler D, Grady C: What makes clinical research ethical? *JAMA* 2000;283:2701-2711.