Pergamon

*European Journal of Cancer* Vol. 33, No. 7, pp. 987–988, 1997
© 1997 Elsevier Science Ltd. All rights reserved
Printed in Great Britain
0959-8049/97 $17.00 + 0.00

**PII: S0959-8049(97)00109-3**

# Editorial

# Evidence Based Medicine, a New Challenge

A. Costa[1] and S.M. Hubbard[2]

[1]European Institute of Oncology, Via Ripamonti 435, 20141 Milan, Italy; and [2]International Cancer Information Center, NCI, Bethesda, Maryland 20892, U.S.A.

DIVERSE efforts have been made in recent years aimed at reducing and/or eliminating gaps between practitioners' knowledge and the most recent clinically relevant scientific information constantly generated by biomedical research. This has led to an increased awareness of the fact that doctors are confronted with a growing body of information, much of it invalid or irrelevant to clinical practice [1].

Interpreting and summarising the literature to generate recommendations for clinical practice thus becomes an important part of the process aimed at health care improvement. However, even when recommendations come from rigorous approaches, it is important to differentiate between those based on inadequate evidence (which might require reversal when sufficient data become available) and those based on strong evidence (which actually can save lives) [2]. The concept of 'evidence based medicine' has recently emerged as a type of medicine able to combine clinical skill with objective evidence of efficacy.

Evidence based medicine (EBM) de-emphasises intuition, unsystematic clinical experience and pathophysiological rationale as sufficient grounds for clinical decision making, and stresses the examination of evidence from clinical research. EBM requires new skills from the physician, including efficient literature searching and the application of formal rules of evaluating the clinical literature [3].

In the field of oncology, in 1982, the U.S. National Cancer Institute (NCI) began a major effort to disseminate advances in cancer treatment through computer technology by introducing the Physician Data Query database (PDQ). PDQ currently contains peer-reviewed monthly updates for over 80 different types of cancer, as well as supportive care statements, screening and prevention information, and summarises descriptions of ongoing and closed clinical trials from around the world. As recently stated by its editor in chief, Daniel G. Heller, PDQ's objective is to provide high-quality interpretation of the literature without bias [4]. It is also worth noting that PDQ is one of the few sources of cancer information not influenced by industry.

Besides the problem of the rapid transfer of knowledge from research to practice, two other issues were addressed in the 1980s by several health care researchers: the cost implications of the medical decision and its quality of life consequences for the patient. The complexity of the decision making process in modern medicine (including, of course, oncology) is thus strictly related to the need to combine efficacy (i.e. maximum chances of survival) with the best possible quality of life together with cost effectiveness.

Two options are currently offered to the clinician willing to cope with this demand for effectiveness: one is more individualised and is represented by already selected sources of information in which all that is considered best practice for that specific clinical situation can be found. This is the case of the already mentioned PDQ database, periodically reviewed by committed experts: the enquiring physician has only to know the right question and he will get the right answer. As shown by the paper of C. Arrigo in this issue (pages 997–1001), PDQ has become one of the most frequently used databases of the U.S. National Library of Medicine. Its use outside the U.S.A. can be limited by the requirement of special equipment and by language barriers, but there is little doubt that this source of information can be of great help to the practicing physician in cancer care.

The second option, less individualised but including a cost-effective evaluation, is the one of the so-called 'clinical practice guidelines' in which the decision making process is made easier by following the recommendations that an external authoritative body has generated for that clinical situation. This is the case, for example, of the SOR database (Standards, Options et Recommandations) produced in France, in the French language, by the Federation of the Anti-Cancer under the guidance of Philip and associates [5].

All guidelines have to face the problem of 'quantifying' the evidence used to determine the guidelines, and for this reason they all rely on a careful evaluation of the literature and of the different sources of published data. Attempts to classify the level of evidence of medical statements date back to the late 1970s in Canada, but increasing attention has been devoted to this problem in recent years. In 1992, the U.S. agency for Health Care Policy proposed a five-level scale of evidence, ranking first 'well-designed' randomised clinical trials or meta-analyses and fifth case reports (Table 1); in the same year the Canadian Task Force on the Periodic Health Examination proposed a scale in which 'at least one properly designed ran-

---

Correspondence to A. Costa.
Received 20 Feb. 1997; accepted 22 Feb. 1997.

*Table 1. Levels of Evidence*

| Level | Evidence |
|---|---|
| I | High power randomised trial or meta-analyis of well-designed randomised clinical trials |
| II | At least one well-designed non-randomised study or one low-powered randomised clinical |
| III | Well-designed quasi-experimental studies (cohort, case-control series, etc.) |
| IV | Well-designed non-experimental studies (comparative, correlational, descriptive, etc.) |
| V | Case reports |

Ag Health Care Pol 1992.

domised trial' is ranked as the first level of evidence, but also 'opinions of respected authorities, descriptive studies and expert committees' are introduced as levels of evidence. Agreement among experts, in the lack of evidence (the so-called 'grey zones' 6 or 7), is also present in the scale used by SOR. In an effort to take into account not only statistical evidence but also clinical relevance, the U.K. EBM working group introduced in 1995 the additional parameter of 'number of patients needed to treat' (NNT) to avoid an undersized event. This approach stresses the importance, not only of the type of studies on which to base the recommendation, but also the expected degree of effect of the recommended decision.

A fifth category on which to base recommendations after the style of study, study design, expert opinions and effect degree is the so-called 'logical inference' from existing evidence, originally proposed by the authors of the database START. START stands for State of the Art in Oncology in Europe and is a hypertext system produced by an active group of Italian medical oncologists (P. Casali, L. Licitra, C. Tondini at the Instituto Nazionale per lo Studio e la Cura dei Tumori, and F. De Braud at the European Institute of Oncology, both in Milan), with the coordination of A. Santoro and F. Cavalli, and under the auspices of the European School of Oncology. START is in English, available on the Internet and represents the result of a long process of consensus among European oncologists (approximately 300 of them are actively involved in this effort).

The originality of the START classification of levels of evidence (Table 2) is exemplified by the innovative concept of 'logical inference' as an indirect basis for treatment recommendations. It is suggested that, sometimes, different pieces of information can be rationally combined with each other and thus justify clinical decisions unsupported in themselves by direct evidence. In this regard, according to the START database, some treatment recommendations can be considered as 'appropriate for individual clinical use', even though they cannot be viewed as proper 'standard' options. An illuminating example is offered by the issue of prophylactic cranial irradiation of patients with stage II small cell lung cancer: it does not alter survival but it does reduce CNS metastases. Why shouldn't it be recommended in specific environments with RT facilities and cost-effectiveness flexibility?

*Table 2. Bases for START recommendations*

| Basis | Evidence |
|---|---|
| C | General consensus |
| 1 | More than one consistent randomised trial and/or meta-analysis |
| 2 | One or a few unconfirmed randomised trial(s) or conflicting results |
| 3 | Non-randomised evidence |
| R | Logical inference from existing evidence |

*Table 3. Levels of evidence for the PDQ*

1. Randomised controlled clinical trial(s)
    i. Double-blinded
    ii. Non-blinded
2. Non-randomised controlled clinical trial(s)
    i. (e.g., allocation by birthdate, chart number, etc.)
3. Case series
    i. Population-based, consecutive cases
    ii. Consecutive cases (non-population-based)
    iii. Non-consecutive cases

Endpoints
A. Total mortality
B. Cause-specific mortality
C. Carefully assessed quality of life
D. Indirect surrogates
    i. Disease-free survival
    ii. Progression-free survival
    iii. Tumour response rate

Developed by B. Kramer and S. Boyer, 10/94.

In conclusion, today's physician needs careful help in coping with the fast pace of data proliferation in the literature. This support should be inexpensive or possibly free and easily accessible. Databases on the Internet may be an answer to this demand, even in countries with limited resources. When offering the physician more than just a selection of the literature, medical information providers enter the field of state-of-the-art medicine (with an implicit labelling of gold standard to their statements) or, moreover, of recommendations and guidelines. The level of evidence of the statements thus becomes essential to give authority to what is recommended. Evidence based medicine sums up the results of available studies and quantifies the strength of medical hypotheses.

A closer comparison of the different scales of levels of evidence (a new one is being adopted by the PDQ database, Table 3) is now mandatory to harmonise the different procedures and processes leading to medical recommendations and guidelines.

---

1. Rosenberg W, Donald A. Evidence based medicine: an approach to clinical problem-solving. *Br Med J* 1995, **310**, 1122–1126.
2. Guyatt GH, Sackett DL, Sinclair JC, Hayward R, Cook DJ, Cook RJ for the Evidence-Based Medicine Working Group. IX. A method for grading health care recommendations. Users' guides to medical literature. *JAMA* 1995, **274**, 1800–1804.
3. Evidence-Based Medicine Group. A new approach to teaching the practice of medicine. *JAMA* 1992, **268**, 2420–2425.
4. Sarosy GA. PDQ Board Visits Europe. *Protocall—The Newsletter of the NCI Information Associates Program.* 2 **(2)**, 1996.
5. Arrigo C. An evaluation of the NCI Physician Data Query (PDQ) Dissemination in Europe. *Eur J Cancer*, **33** (7), 997–1001.
6. Naylor CD. Grey zones of clinical practice: some limits to evidence-based medicine. *The Lancet* 1995, **345**, 840–842.

**Acknowledgement**—The authors of this editorial are much obliged to the START team for having provided the relevant information on the different classifications of levels of evidence.