Are Therapeutic Decisions Supported by Evidence From Health Care Research?

Gaétane Michaud, BS, BEd, MD; Jessie L. McGowan, BMus, MLIS; Richard van der Jagt, MD, FRCPC; George Wells, PhD; Peter Tugwell, MD, MSc, FRCPC

Background: One of the most common decisions physicians face is deciding which therapeutic intervention is the most appropriate for their patients. In recent years much emphasis has been placed on making clinical decisions that are based on evidence from the medical literature. Despite the emphasis on incorporation of evidence-based medicine into the undergraduate curriculum and postgraduate medical training programs, there has been controversy regarding the proportion of interventions that are supported by health care research.

Objective: To investigate the proportion of major therapeutic interventions at our institution that are justified by published evidence.

Methods: One hundred fifty charts from the internal medicine department were reviewed retrospectively. The main diagnosis, therapy provided, and patient profile were identified and a literature search using MEDLINE was performed. A standardized search strategy was developed with high sensitivity and specificity for identifying publication quality. The level of evidence to support each clinical decision was ranked according to a predetermined classification. In this system there were 6 distinct levels, which are explained in the study.

Results: Of the decisions studied, 20.9% could be supported by placebo-controlled randomized trials and 43.9% by head-to-head trials. Half of these were shown to be significantly superior to the treatment against which it was being compared. For 10 of the 150 clinical decisions, evidence was found demonstrating alternative therapies as being more effective than that selected.

Conclusions: Most primary therapeutic clinical decisions in 3 general medicine services are supported by evidence from randomized controlled trials. This should be reassuring to those who are concerned about the extent to which clinical medicine is based on empirical evidence. This finding has potential for quality assurance, as exemplified by the discovery that a literature search could have potentially improved these decisions in some cases.

Arch Intern Med. 1998;158:1665-1668
METHODS

Our study used a retrospective review of a systematic sample of charts of patients admitted to the Clinical Teaching Unit (CTU) at the Ottawa General Hospital, a university hospital in Ottawa, a mid-sized Canadian city. The CTU consists of 3 general medicine teams, each composed of a staff member, a senior resident, 1 or 2 junior residents, and 1 or more senior medical students. Patients’ charts for all admissions on 2 preselected days per month for a 1-year period (June 1994–June 1995) were reviewed. For any patient readmitted with the same diagnosis, only the initial hospital admission was included within the study results. A standardized data abstraction form was developed and used to record (1) the most responsible diagnosis, (2) the major therapeutic maneuver (extracted from the discharge summary as dictated by the physician directly caring for the patient), (3) patient population characteristics, and (4) the desired outcome. Disease categories were defined by subcategories as outlined in the International Classification of Diseases, Ninth Revision (ICD-9). The OVID 3.0 version of MEDLINE from January 1966 to June 1995 was used to conduct the search for the best evidence, using a modification of the search strategy from Haynes et al; the search started with the most recent databases and was discontinued once all databases were searched or an article was retrieved containing a statistically and/or clinically significant placebo-controlled RCT supporting the use of the intervention used by the team. A checklist was developed that captured the key methodological elements recommended by Guyatt et al and by Sackett et al. This differed from the scale used by Ellis et al in 2 respects: first, the information collected included an assessment of the statistical power (P≤.05) to ensure that the trials were sufficiently large to draw a valid conclusion; second, we distinguished between observational study designs that had control groups (cohort or case-control designs) from simple case series without controls. The articles are classified based on the quality level of the evidence. Levels are the following:

- Statistically significant placebo-RCTs (low α error, P≤.05)
- Nonsignificant placebo-RCTs
- Statistically significant head-to-head trials (low α error, P≤.05)
- Nonsignificant head-to-head trials
- Case-control and cohort studies
- Case series without control subjects
- No supporting evidence retrieved

Primary analysis comprised descriptive statistics of disease classification and related therapy(ies). With a sample size of 150 patients, and a hypothesized estimate of 20% of clinical decisions being based on statistically and clinically significant placebo-RCTs, the 95% confidence interval ranges from 14% to 26%.

To compare the results obtained at our institution with those from the study by Ellis et al, the charts were then reviewed and classified according to the 3 following levels: (1) interventions without substantial evidence, (2) convincing nonexperimental evidence, and (3) evidence from RCTs. Physician experts in the appropriate subspecialty who were blinded to the chart review and its hypothesis were charged with reviewing each case and classifying it according to the above-mentioned classification.

RESULTS

Of 150 charts selected for review, 148 met the inclusion criteria. The remaining 2 cases were patients already enrolled in the study and were consequently readmitted for the same diagnosis. To demonstrate the generalizability of the types of questions posed at our center, the decisions for which statistically significant placebo-RCTs were retrieved are presented as a sample in the Table. The data collected were categorized according to the expanded classification schema described earlier. These results are presented in Figure 1.

Of 148 cases analyzed, 31 of the clinical decisions could be supported by a placebo-RCT demonstrating a clinically important and statistically significant (P≤.05) benefit over placebo. The remaining 3 clinical trials showed no statistical benefit over placebo for that specific diagnosis even though the statistical power was adequate.

In 65 of the cases the major intervention, although not tested in a placebo-controlled trial, was supported by a head-to-head trial (65 [43.9%] of the references retrieved). Thirty-three of these showed one intervention...
to be statistically superior over the therapy to which it was being compared and the drug of choice was used in all 33 of these cases. In the remaining 32 cases no statistical advantage was demonstrated between the 2 treatment arms. In 28 of these cases the therapy chosen by the CTU was the standard therapy that constituted the control arm in these studies. Case-control or cohort studies included 23 (15.5%) of these studies. Case series accounted for 4 (2.7%) of those located. Cases for which no study was found to support the intervention accounted for the remaining 22 (14.7%).

In 10 cases (6.7%) an alternate therapy was identified from a controlled trial that was clinically and statistically superior to that chosen. This subgroup includes cases from all levels. Examples from this group include (1) giving omeprazole alone for peptic ulcer disease; (2) prescribing pyrazinamide alone for tuberculosis; and (3) using amino glutethimide for Cushing disease rather than ketoconazole.

As stated earlier, to allow for comparison of results with the study by Ellis et al., our group classified the results of this study according to their classification. These results are given in Figure 2. As may be seen, the results were similar to those obtained by Ellis and colleagues. In 57% of cases in this study, the evidence to support the decisions made could be supported by RCTs, either placebo controlled or head to head, compared with 53% in the study by Ellis et al. Our group classified 27% of the decisions made by our CTU as convincing nonexperimental evidence. This result was also in keeping with the results of 29% by Ellis et al. Finally, interventions without substantial evidence were found to encompass 16% of CTU decisions in this study, compared with 18% in the study by Ellis et al.

The results of this study demonstrate that most clinical decisions made in the CTUs at our institution were supported by the best available evidence from health care research and positively support those found by Ellis et al. The robustness of this result is strengthened because this involved assessment of the best available evidence to support therapy decisions made by the 24 different teams studied. In addition, it aids to refute the misperception that a significant proportion of clinical decisions made by physicians may be erroneous and/or hazardous.

The generic ability of this result is limited by the fact that it was carried out at an internal medicine teaching service in a university hospital setting. Thus, this study needs to be repeated in different disciplines and settings before generalizing the results to other groups and sites. However, at the hospital where this study was performed, the CTU admits more than 95% of its patients from the emergency department that provides care to the local community; most patients are admitted to internal medicine and the patients admitted have diagnoses common to the practice of internal medicine anywhere in the world, as can be seen from the Table.

Evidence-based medicine has limitations like most things in medicine. There exist ethical problems conducting placebo-controlled long-term studies; large sample sizes are needed to detect rare adverse effects and the effects of comorbidities cannot be assessed in most trials. We acknowledge that there remains a necessity to obtain nonexperimental studies in the absence of RCTs in the setting of decisions dealing with rare conditions and those not ethically amenable to RCT methodology.

However, there is consensus on the use of being able to show that a clinical decision is supported by documented benefit as demonstrated in a methodologically

---

**Diagnoses and Therapies Provided for Which Placebo–Randomized Controlled Trials Existed**

<table>
<thead>
<tr>
<th>Diagnosis</th>
<th>Therapy</th>
<th>No.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Angina</td>
<td>Aspirin</td>
<td>1</td>
</tr>
<tr>
<td>Asthma exacerbation</td>
<td>Methylprednisolone</td>
<td>7</td>
</tr>
<tr>
<td>Asthma exacerbation</td>
<td>Prednisone</td>
<td>3</td>
</tr>
<tr>
<td>Cerebrovascular accident</td>
<td>Aspirin</td>
<td>1</td>
</tr>
<tr>
<td>Congestive heart failure</td>
<td>Captopril</td>
<td>4</td>
</tr>
<tr>
<td>Congestive heart failure</td>
<td>Furosemide</td>
<td>4</td>
</tr>
<tr>
<td>Cystic fibrosis exacerbation</td>
<td>Tobramycin aerosol</td>
<td>4</td>
</tr>
<tr>
<td>Depression</td>
<td>Fluoxetine</td>
<td>1</td>
</tr>
<tr>
<td>Epileptic seizure</td>
<td>Carbamazepine</td>
<td>2</td>
</tr>
<tr>
<td>Epileptic seizure</td>
<td>Phenytoin</td>
<td>2</td>
</tr>
<tr>
<td>Herpes simplex virus infection</td>
<td>Acyclovir</td>
<td>1</td>
</tr>
<tr>
<td>Peptic ulcer disease</td>
<td>Amoxicillin</td>
<td>1</td>
</tr>
</tbody>
</table>

---

**Figure 1. Evidence to support clinical decisions. Stat indicates statistically; sig, significant; nonsig, nonsignificant; and RCT, randomized controlled trial.**

**Figure 2. The results of Ellis et al. and ours at Ottawa General Hospital (OGH), Ottawa, Ontario, according to the classification of Ellis et al. RCTs indicates randomized controlled trials.**

---

©1998 American Medical Association. All rights reserved.
sound study. Although we accept that “not all that is measured is of value, and not all that is valued can be measured,” critical appraisal skills are essential if physicians are to be able to interpret and integrate the scientific basis of medicine where it exists. The ability to detect clinically important and statistically sound studies is a skill that can be acquired by clinicians who are not trained to carry out research themselves. In this age of increasing consumerism, patients have a right to expect their clinician to provide them with quantitative estimates of benefit and risk. Critical appraisal skills are fundamental if the evidence is to be presented in a balanced fashion, using data sources such as the Cochrane Collaboration Library.

Our study also has application in the area of quality assurance. We believe that quality assurance in the provision of optimal care to our patients is provided by ensuring that our clinical decisions have been informed by health care research whenever possible. The fact that such a large proportion of care studied is indeed justifiable on the basis of sound research is a powerful argument for its incorporation into quality assurance programs. The finding that a literature search could have potentially improved on these decisions in some cases further supports its justification as a powerful tool in quality assurance programs. Furthermore, it can help clinicians avoid being pressed by their patients into making clinical decisions based on the latest media headline or Web site on the Internet, which the patient has obtained from the now widely accessible consumer health information. As there are no means of quality assurance for these sources of information, they may evoke clinical decisions made without foundation. The avoidance of decisions based on unfounded benefit may be potentially health and cost saving, and therefore should appeal to those individuals or establishments paying for health care, clinicians, and the general public. Such an assurance strategy also links the evaluation of health care to the curriculum and role modeling by the teachers in university training institutions. Health outcomes are certainly the final truth, and can be usefully applied as a trigger to examine conditions in detail where the outcomes vary excessively in internal medicine from local or national norms.

In conclusion, the widespread use of EBM, a scientific approach to clinical decision making, as a component of clinical practice may be of benefit to both the clinician and his/her patient. However, to date the exact value of this method on clinical decision making is yet unmeasured. The ultimate success of EBM as a therapeutic tool is contingent on continuous faculty and student development. This priority has been identified in initiatives such as the Educating Future Physicians for Ontario (EFPO) Project and the CANMEDS 2000 Project of the Royal College of Physicians and Surgeons of Canada.

Accepted for publication January 8, 1998.

Corresponding author: Peter Tugwell, MD, MSc, FRCPC, Department of Medicine, University of Ottawa, 501 Smyth Rd, Ottawa, Ontario, Canada K1H 8L6.

REFERENCES

9. Guyatt G, Sackett D, Cook D, for the Evidence-Based Medicine Working Group. Users’ guides to the medical literature, II: how to use an article about therapy or prevention, B—what were the results and will they help me in caring for my patients? JAMA. 1994;271:59-63.