Evidence-based surgery

Mohey Elbanna, MD
Department of General Surgery, Ain Shams University

In the light of the national focus on patient safety, understanding the principles of evidence-based surgery and its implications in surgical practice has gained great importance. Evidence-based surgery is even more important in the face of regular attacks on surgeons, by scientists, for the perceived deficiencies of the evidence base in their specialty. Additionally, there is rising demand for surgical services, plan for universal health insurance, emergence and spread of private health insurance companies, patients' increasing dependence in selecting their own care, and the widespread media coverage of errors and alleged errors paralleled with increased legal disputes over malpractice.

All these and other factors, as the natural eagerness of each surgeon to improve her/his own practice, require all surgeons to understand the development and application of evidence-based surgery and surgical outcomes research.

An analysis of general surgical work in a large UK hospital showed that only 24% of the treatments used were based on randomized controlled trials (RCT) evidence, compared with over 50% for inpatient general medicine. A recent analysis of the illnesses and treatments most commonly encountered in general surgery suggested that less than 40% of operative treatments were amenable to study using an RCT design. Some of the reasons suggested for this, such as the rarity or emergency nature of the conditions involved, are not wholly convincing, but others are important.

What is evidence-based surgery?

The term Evidence-based medicine is the one typically used in the literature. It is applicable to surgery and other fields as well. There are several definitions of Evidence-based medicine. Following are two of the most valuable. Evidence-based medicine has been variably defined as: the conscientious, explicit and judicious use of current best evidence in making decisions about the care of individual patients and the integration of individual clinical expertise with the best available external clinical evidence from systematic research.

Alternatively, Evidence-based medicine is defined as an information and learning strategy that seeks to integrate clinical expertise with best evidence available in order to make effective clinical decisions that will ultimately improve patient care.

Framework for evidence-based surgery:

Surgical outcomes research studies the quality of surgical care, from the selected treatment approaches and clinical interventions to systems of care delivery. The results are usually of interest to physicians, patients, as well as policymakers. Key considerations in evidence-based surgery include the measurement and interpretation of outcomes and the design of studies to obtain them. The outcomes of healthcare are measured by examining the structures and processes of care as well as patients' factors and risks. Structural features comprise patient, provider, and payer characteristics. Process-of-care measures describe what was done to and for the patient. Outcomes are then classified as clinical and physiologic outcomes, patient- reported outcomes and economic outcomes.

Clinical outcome measurement:

The most commonly used clinical measures of the consequences of surgical therapy are rates of mortality and morbidity, gains in life expectancy, relative risk, relative and absolute risk reduction, and potential impact in absolute numbers.

Mortality is the most reliably measured clinical outcome. It is expressed as the proportion of deaths from a particular cause over a defined time interval and most reliably measured from death certificates. In mortality rates, the denominator represents the entire population at risk of dying from the disease. Case fatality rate includes only those with the

disease in the denominator. However, mortality has less meaning in the study of surgical procedures in which death is rare; therefore, it is more meaningful to report morbidity, often referred to in the surgical literature as complications.

Economic outcomes measurement:

Concerns about the costs of healthcare have dramatically increased the demand for economic outcome measures. They quantify the costs and/or benefits of medical and surgical care. In evidence-based surgery literature, clinical measures of outcome such as mortality or treatment complications are reported much more frequently than are economic measures, either alone or combination with clinical measures. This reflects interests of investigators as well as difficulties in measuring economic outcomes. Economic measures for clinical outcome studies may be developed in several ways. Cost analysis can be completed prospectively as part of a clinical trial. Retrospective cost data can be analyzed typically based on secondary or administrative database. Typically reported costs include hospital and physician charges, obtained from billing data.

Patient-reported outcomes measures:

Health status, functional status, quality of life, and health-related quality of life are terms used almost interchangeably to refer to the concept of patient reports of their own health. These data are usually collected using standardized questionnaires or surveys. Most surgical studies before 5 years neglected to collect standardized data about patient-reported health status and quality of life.

Quality-of-life assessment can have relevance to surgical research and practice for defining the indication for surgery, for monitoring the patient, and for evaluating the impact of treatment. Surgical quality-of-life assessments should be made when different treatment alternatives might affect the patient's quality of life differently, when new interventions are implemented, when there is scarcity of resources, when the timing of an operative intervention must be determined, and when improving the quality of life is the goal of intervention. When surgery is clearly life-saving or is the only treatment alternative, quality of life assessment may be less important.

Levels of evidence:

Certain fundamentals apply to the analysis of the available data from surgical practice. These are best considered in the context of the evidence-based surgery movement. The traditional levels of scientific investigation are presented in , from the exalted RCT to the lowly case report and expert opinion.

Table (1): Levels of Evidence-Based Medicine (Institute for Clinical Systems Improvement).

Level	Study Type
I	Randomized Clinical Trial
II	Retrospective Review
III	Nonconsecutive small population study
IV	Case Report
V	Expert Opinion

Randomized Controlled trials are not the start point of scientific evidence. Actually, it is the endpoint. The conscientious surgical investigator must grow the idea on a foundation of expertise (expert opinion, level V). Then, he is to try it once (case report, level IV). If the idea works, he should try it a couple more times (nonconsecutive, small population study, level III). At this point, he may think he is on to something, so he scrounges around collecting anyone else's similar experience (retrospective review, level II). If he has been both persistent

and lucky, he talks to colleagues and patients into an RCT (level I). This means that all levels of scientific inquiry enjoy an important place in the sequence of investigation leading to an advance in surgical therapy. Moreover, the pyramids of surgical practice and surgical science look contrary to each other. However, this has a historical meaning and a practical meaning. For well established data and treatments, surgical practice should be based on evidence concluded from RCTs, i.e. from the top of the pyramid of evidence. For new fields of diagnosis and treatment, the contrary is true. We start from the base - not the bottom -up.

There are different grades of evidence. Here the one suggested by the Institute for Clinical Systems Improvement (ICSI) is presented.

Grades of evidence:

Grade I: The evidence consists of results from studies of strong design for answering the question addressed. The results are both clinically important and consistent with minor exceptions at most. The results are free of any significant doubts about generalizability, bias, and flaws in research design. Studies with negative results have sufficiently large samples to have adequate statistical power.

Grade II: The evidence consists of results from studies of strong design for answering the question addressed, but there is some uncertainty attached to the conclusion because of inconsistencies among the results from the studies or because of minor doubts about generalizability, bias, research design flaws, or adequacy of sample size. Alternatively, the evidence consists solely of results from weaker designs for the question addressed, but the results have been confirmed in separate studies and are consistent with minor exceptions at most.

Grade III: The evidence consists of results from studies of strong design for answering the question addressed, but there is substantial uncertainty attached to the conclusion because of inconsistencies among the results of different studies or because of serious doubts about generalizability, bias, design flaws, or adequacy of sample size. Alternatively, the evidence consists solely of results from a limited number of studies of weak design for answering the question addressed.

Grade Not Assignable: There is no evidence available that directly supports or refutes the conclusion.

Classes of research reports:

The Institute for Clinical Systems Improvement (ICSI) has suggested the classes of research reports, presented in **Table(2)**.

Table (2): Classes of research reports (ICSI).

A. Primary Reports of New Data Collection:
Class A: Randomized, controlled trial
Class B: Cohort study
Class C: Non-randomized trial with concurrent or historical controls Case-control study Study of sensitivity and specificity of a diagnostic test Population-based descriptive study
Class D: Cross-sectional study Case series Case report

B. Reports that Synthesize/Reflect on Collections of Primary Reports:
Class M: Meta-analysis Systematic review Decision analysis Cost-effectiveness analysis
Class R: Consensus statement Consensus report Narrative review
Class X: Medical opinion

Study design:

The quality of the evidence and the strength of the study design are critical factors in conducting and evaluating outcomes research. Determining causal links between two variables (to establish which procedures are more effective) requires randomized clinical trials or matched-pair experimental studies with blinding. Randomized clinical trials are generally accepted as the definitive approach for assessing the efficacy of a new treatment. The process of randomization, when properly implemented, provides the means by which the factors that may influence the results of a trial are equally distributed between the experimental treatment group and the control group. The costs of such studies are generally considerable.

It is impossible to subject all new therapies to a randomized clinical trial evaluation, in part because of those costs. Moreover, these studies take time to obtain the outcomes, and in some instances, undertaking a randomized trial is simply not ethical - for example withholding appendectomy for acute appendicitis to determine whether antibiotic treatment alone would be efficacious. In surgery it is even more difficult to perform RCTs, as many surgeries can not be experimentally tried on patients. In addition, RCT may not be based on random samples of patients. The investigators may seek to exclude all but a subset of patients with a particular disease. It is therefore often difficult to generalize the results to the population of patients with that particular disease. These weaknesses must be carefully weighed against the considerable strengths of this type of study.

Meta-analysis:

A useful definition of meta-analysis was given by Huque: "A statistical analysis that combines or integrates the results of several independent clinical trials considered by the analyst to be 'combinable.'" In , a meta-analysis combines the results of several studies that address a set of related research hypotheses. Meta-analysis is performed to overcome the problem of reduced in studies with small sample sizes; analyzing the results from a group of studies can allow more accurate data analysis. Meta-analysis assembles existing research findings to provide an aggregate view. In meta-analysis, the investigator reviews the literature for all relevant studies regarding a given surgical procedure and a specific outcome.

Meta-analysis has several advantages. First, well conducted meta-analysis allows for a more objective appraisal of the evidence, which may lead to resolution of uncertainty and disagreement. Secondly, meta-analysis may reduce the probability of false negative results and thus prevent undue delays in the introduction of effective treatments into clinical practice. Thirdly, meta-analysis of a large number of individual studies or of individual patient data allows testing of a hypothesis regarding treatment effects in subgroups of patients. Fourthly, Heterogeneity between study

results may be explored and sometimes explained. Lastly, promising research questions to be addressed in future studies may be generated, and the sample size needed in future studies may be calculated accurately

However, meta-analysis is based on the assumption that the quality of the individual studies are the same, that the factors examined in the studies are the same, that the data missing from any one study will not be biased for the outcomes of interest, that the population from which the study subjects were drawn are similar, and that the definitions used among the studies are the same. However, it is known that there is "publication bias." Publication bias means studies that do not attain statistical significance are not published as frequently as those that do. Consequently, meta-analysis has its weaknesses.

Because all evidence gathering studies have their weaknesses, good doctors use both individual clinical expertise and the best available external evidence, and neither alone is enough. Without clinical expertise, practice risks becoming tyrannized by evidence, for even excellent external evidence may be inapplicable to or inappropriate for an individual patient. Without current best evidence, practice risks becoming rapidly out of date, to the detriment of patients.

Evidence based surgery is not restricted to randomized trials and meta-analyses. It involves tracking down the best external evidence with which to answer our clinical questions. To find out about the accuracy of a diagnostic test, we need to find proper cross sectional studies of patients clinically suspected of harboring the relevant disorder, not a randomized trial. For a question about prognosis, we need proper follow up studies of patients assembled at a uniform, early point in the clinical course of their disease. And sometimes the evidence we need will come from the basic sciences such as genetics or immunology. It is when asking questions about therapy that we should try to avoid the non-experimental approaches, since these routinely lead to false positive conclusions about efficacy. Because the randomized trial, and especially the systematic review of several randomized trials, is so much more likely to inform us and so much less likely to mislead us, it has become the "gold standard" for judging whether a treatment does more good than harm. However, some questions about therapy do not require randomized trials (successful interventions for otherwise fatal conditions) or cannot wait for the trials to be conducted. And if no randomized trial has been carried out for our patient's predicament, we must follow the trail to the next best external evidence and work from there.

Evidence based medicine is not "cookbook" medicine. Because it requires a bottom up approach that integrates the best external evidence with individual clinical expertise and patients' choice, it cannot result in cookbook approaches to individual patient care. External clinical evidence can inform, but can never replace, individual clinical expertise, and it is this expertise that decides whether the external evidence applies to the individual patient at all and, if so, how it should be integrated into a clinical decision. Similarly, any external guideline must be integrated with individual clinical expertise in deciding whether and how it matches the patient's clinical state, predicament, and preferences, and thus whether it should be applied. Clinicians who fear top down cookbooks will find the advocates of evidence based medicine joining them at the barricades.

Despite its ancient origins, evidence based medicine remains a relatively young discipline whose positive impacts are just beginning to be validated, and it will continue to evolve. This evolution will be enhanced as several undergraduate, postgraduate, and continuing medical education programs adopt and adapt it to their learners' needs. These programs, and their evaluation, will provide further information and understanding about what evidence based medicine is and is not.

References:

1- Sackett DL, Rosenberg WMC, Muir Gray JA, Haynes RB, Richardson WS: Evidence based medicine: what it is and what it isn't. *BMJ* 1996; 312: 71-72.

2- Landry MD, Sibbald WJ: From data to evidence: evaluative methods in evidence-based medicine. *Respir Care* 2001; 46: 1226-1235.

3- Institute for Clinical Systems Improvement (ICSI): Preoperative evaluation. Bloomington (MN): *Institute for Clinical Systems Improvement (ICSI)* 2006; p.33.
4- Arya J, Wolford H, Harken AH.: Evidence-based science; a worthwhile mode of surgical inquiry. *Arch Surg*. 2002; 137:1301-1303.
5- Davidoff F, Haynes RB, Sackett DL, Smith R: Evidence based medicine. *BMJ*. 1995; 310:1085-1086.
6- McCulloch P: Evidence-Based General Surgery. Walton & Fazakerly Hospital University of Liverpool.
7- Egger M, Smith G D: Meta-analysis: Potentials and promise. BMJ No 7119 Volume 315.
8- Gordon T A: Evidence-Based Surgery. In: Greenfield's surgery; scientific principles and practice; 4th edn; 2006; Chapter 14: 318-326.