Evidence-based surgery

Alexander D Shearman

Clifford P Shearman

Abstract

Every year an estimated 234 million major surgical procedures are undertaken worldwide. In 2009–10, 4.8 million hospital admissions involved surgical input in England alone, and around 4 in 5 adults are likely to have an operation in their lifetime. Despite these enormous numbers, lack of objective evidence for the indications and benefits (or otherwise) of surgical procedures is often lacking. Lack of robust research into surgical disease and treatments has been criticized. Less than 5% of national funding for health research involves surgery. This seems surprising as inappropriate surgical treatments can be hazardous for the patient and costly to the health care system. The demand for evidence-based clinical practice is increasing, driven by public and professional expectations. The scarcity of high-quality studies across many different fields of surgery has led to ambiguity in the management of many common surgical conditions with widely varying clinical outcomes in different geographical areas. Surgical treatments are costly and need to be justified not only on clinical benefit, but on their cost effectiveness compared to other treatments. Several approaches have been adopted to evaluate evidence of benefit for surgical treatments. This article outlines these and their application in a clinical setting. The components of evidence-based medicine and the GRADE method of evaluating quality of evidence are explored. The importance of taking into consideration cost effectiveness and patient attitudes to treatment are also discussed.

Keywords Evidence-based medicine; meta-analyses; randomized controlled trials; surgical research; systematic review

Background

I think your solution is just; but why think, why not try the experiment

John Hunter FRS FRCS (1728–1794).

John Hunter made important contributions to surgical science by carrying out experiments and observations on both animals and humans. It has been argued that surgeons have subsequently failed to adopt modern and more appropriate research tools such as randomized controlled trials (RCTs), falling behind other specialties in evaluating treatments for their patients.[1]

Wound care is a good example of an area where there is vast expenditure on complex dressings, but little, if any, high-quality evidence for their role.[2] Not only is lack of evidence frustrating when planning treatment, it can potentially damage or prevent the adoption of an effective therapy. Surgeons managing people with complications of diabetic foot disease have enthusiastically adopted topical negative pressure wound therapy (NPWT). Early, smaller studies have shown this treatment to be expensive, possibly associated with significant complications and delays in discharge from hospital, despite possible benefits in initial healing rates. Purchasers of health care are increasingly looking for evidence of clinical and cost effectiveness and without it are unlikely to support the treatment.[3] It is only very recently that well-constructed systematic reviews have shown NPWT to be advantageous over other treatments, both in terms of clinical efficacy and cost.[4]

On the other hand, some of the most complex surgical treatments have been successfully studied and current practice is now based on the study findings. In 1985 over 100,000 carotid endarterectomies were performed in the United States alone costing over 1 billion dollars. It was estimated that a third of these cases were inappropriate and a further third were of dubious benefit. The European Carotid Surgery Trial (ECST) and the North American Symptomatic Carotid Endarterectomy Trial (NASCET) published their results in 1991. Based on a combined total of 3730 enrolled patients the studies clearly demonstrated benefit, in terms of stroke reduction, in those patients with severe carotid disease offered surgery in combination with best medical treatment when compared to best medical treatment alone. This has allowed clear criteria to be developed to select patients who will gain the most benefit for carotid endarterectomy[5] and the procedure is now recognized as a significant tool in reducing stroke in the National Stroke Strategy.

Evidence-based medicine

Although the concept had been discussed before, the practice of evidence-based medicine (EBM) emerged in the 1980s and 1990s in an attempt to encourage clinicians to critically appraise evidence more thoroughly when treating patients. It represented a generalized dissatisfaction on the reliance of expert opinion and individual approaches to treatments, resulting in a wide range of therapies offered to patients with the same condition. EBM is '*the conscientious, explicit, and judicious use of current best evidence in making decisions about individual patients*'.[6] The principle of EBM is to identify the best available evidence, evaluate it and use it in combination with clinical expertise to aid the treatment of patients. This approach has been successfully utilized by groups resourced to carry out detailed evaluations of therapies which can inform national policy:

- Scottish Intercollegiate Guidelines Network (SIGN)
- National Institute for Health and Care Excellence (NICE).

An evidence-based clinical guideline relieves the individual clinicians of carrying out detailed analysis themselves, uses the most detailed evaluation tools possible and can be regularly updated (https://www.nice.org.uk/guidance). The practice of evidence-based medicine can be divided into four key stages.

The research question to be answered

Having identified the clinical problem that is to be studied, a clear research question must be designed so that current evidence can clearly be identified in the published literature. If the

Alexander D Shearman *BSc MBBS FRCS(Tr&Orth) is a Specialist Registrar in Trauma and Orthopaedic Surgery in the London Deanery, UK. Conflicts of interest: none declared.*

Clifford P Shearman *BSc MBBS FRCS MS is Emeritus Professor of Vascular Surgery at the University of Southampton, UK. Conflicts of interest: none declared.*

Please cite this article in press as: Shearman AD, Shearman CP, Evidence-based surgery, Surgery (2018), https://doi.org/10.1016/j.mpsur.2018.07.010

research question is vague or too broad, then the returns from literature searches will be vast and unhelpful. The PICO structure is commonly used to design the question (Table 1).

Using this approach allows a literature search to be conducted which is likely to identify key information in the field. It is also a useful exercise when reading a research paper to identify exactly what question the authors were trying to answer; it is often is not very clear! A well-constructed hypothesis is central to all high-quality research.

Searching the existing evidence

It is wise to start to look at EBM resources such as NICE and SIGN. If a topic has been covered then the methodology and analysis will be described in detail including evidence that was not included. The Cochrane Library is a collection of six databases of high-quality, independently collected evidence that includes systematic reviews and a registry of randomized controlled trails.[7] To search published evidence keywords from the research question should be entered into the medical subject heading (MeSH) of a search engine such as EMBASE or Medline. Medline can be accessed via PubMed which is hosted by the US National Library of Medicine of the National Institute of Health. EMBASE (Excerpta Medica Database) contains biomedical and pharmaceutical data and is available through NHS Evidence. Depending on the keywords used for the search a number of articles will be identified, which can then be filtered further. Reviewing the abstracts allows irrelevant articles to be discarded and the remainder can then be critically reviewed. A typical search may produce over 2–3000 returns of which approximately 2–30 may be relevant and warrant further analysis.

Critical appraisal of the evidence

There is a hierarchy of evidence values based on the study design (Table 2). Although RCTs are generally held to be the best study design individually, they may have limitations and combining a series of RCTs in a systematic review or combining the data sets from a number of RCTs in a meta-analysis produces more powerful evidence (Box 1). RCTs should be reported according to Consolidated Standards of Reporting Trials (CONSORT) guidelines.[8]

PICO structure

Population (patient or problem)	Population that is of interest, e.g. people with diabetes and foot wounds and arterial disease
Intervention (prognostic factor or exposure)	Intervention that is to be examined, e.g. topical negative pressure therapy after digit amputation
Comparison intervention (if appropriate)	Comparison intervention, e.g. standard healing by secondary intention
Outcome you would like to measure or achieve	Outcome, e.g. complete wound healing

Table 1

A systematic review identifies relevant studies for a specific research question from the literature, assesses the quality of evidence produced and combines the results of those of considered high enough quality. If the raw data are clear, or the authors can help verify it, the results from a series of studies may be combined mathematically to provide a more comprehensive meta-analysis. This has the advantage of increasing the numbers of subjects and hence the power of the study to detect a real difference. However, this approach is based on the assumption that all the study populations are similar, and the methodology used was the same in all the reports included. The Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) checklist, published in 2009, ensures quality and reproducibility when designing these studies.[9]

The results of both can be graphically represented as Forest plots, illustrating the size and the confidence intervals of the effects seen. In the example shown in Figure 1 the number of events in a treatment group compared to those who received placebo in a control group is compared. Studies identified of high quality are listed with the number of events that occurred in each arm of the study. The vertical line passing through 1 indicates no effect. Each study is displayed as a box the size of which represents the sample size. Passing through this the horizontal line represents the 95% confidence intervals. If the box appears on the favours treatment side but the confidence intervals cross the 'no effect' line then there is a chance that treatment is not effective. Pooled results are represented by diamonds of which the right and left extent indicates the confidence. Weighting indicates the influence of the individual study on the pooled results. The odds ratio is a measure of the size of the effect and Mantel-Haensel (M-H) statistic is a method of determining pooled odds ratios.

In a cohort study one group who has received an intervention is compared to another group who did not. There is no randomization and a number of factors such as case selection and surgeon preference for a treatment may confound the results. However, prospective cohort studies can be very important in surgical disease, for example identifying the significance of micro metastasis in lymph nodes removed during breast surgery.[10]

In a case-controlled study a patient group is compared retrospectively to a group matched for a number of variables such as age or sex. A case series is a selected number of patients (not usually consecutive) who have a condition or treatment. There is no comparator group and often the observations reflect the opinions of the authors. Case series however, may be helpful in identifying a particular area that needs further research. Individual case reports have no role in changing practice but may be useful educationally to raise awareness of an unusual condition.

If there is no evidence available on which to develop a clinical guideline then expert opinion can be used. Although this can be very biased, it is possible to extract more value out of the process by challenging groups of experts with a series of clinically relevant questions and debating an agreement; a Delphi Consensus.[11]

Registry data are increasingly being used to inform surgical practice. The National Joint Registry and National Vascular Registry are good examples of this (see '*Evaluating quality in clinical care*' in this issue).

Please cite this article in press as: Shearman AD, Shearman CP, Evidence-based surgery, Surgery (2018), https://doi.org/10.1016/j.mpsur.2018.07.010