

Perspective

Evidence-Based Medicine in Endometriosis Surgery: Double-Blind Randomized Controlled Trial Versus the Consensus Opinion of Experts

Medicine used to be an empirical set of methods and nonempirical principles acquired by observation and research to prevent, diagnose, and treat or palliate the disease and help the patient. Medicine is considered an art because the interpretation of complaints and symptoms and the choice of treatment require special skills and individualization. The state of the art progressively evolves and improves by trial and error guided by the available scientific and observational knowledge. This knowledge is shared freely.

Research has progressively added knowledge and evidence. We tend to forget that available proven evidence constitutes only a small part of decision-making and of selecting medical treatment or surgery in daily practice. Taking hysterectomy as an example, let us consider all the variables that occur: determining the indication for surgery, selection of the surgical access route and technique, selection of sutures and energy, and the many other parameters that vary within the operating theater environment and with the skill and experience of the surgeon. This plethora of variables and the absence of unanimous agreement on many aspects of this process highlight the problem. Nevertheless, in the absence of concrete evidence, the surgeon must use the available knowledge and make decisions based on experience, the practice of the institution, and the regulations of the jurisdiction in which he or she works.

Knowledge of anatomy and understanding the mechanisms involved, as established by research, guides the progressive development of surgery and of the choices involved. Most have become basic surgical principles: preservation of anatomy and function, use of proper dissection planes, and meticulous hemostasis and nerve sparing. The latter was empirically introduced in oncologic surgery; however, it is only over the last decade that we began to understand the anatomy of the pelvic nerves and started to incorporate nerve sparing into reproductive and endometriosis surgery. As surgeons, we almost

unconsciously incorporate into our practice each little piece of evidence that we absorb at meetings and congresses and discussions with our colleagues. After having realized a month ago during a lecture by R. Botchorichvili the risk to traumatize the right hypogastric nerve during promontofixation, it would be impossible not to use this information the next day. Although the physics and tissue effects of various energy modalities used are well known, each surgeon has his or her own preferences, usually based on training as well as the availability and the ability to use the same source of energy without the need of change during the intervention.

Evidence-based medicine (EBM) has been defined as follows: “EBM integrates clinical experience and patient values with the best available research information” [1]. The randomized controlled trial (RCT) and its meta-analysis is obviously at the top of the pyramid of strength of evidence, whereas ideas and opinions are at the bottom. Although not stated explicitly, surgical experience seems to be relegated to the group of “lowest evidence.” Not surprisingly, EBM has become a movement in crisis with the realization that clinical guidelines have become unmanageable, that what is statistically significant may have little clinical relevance, and that EBM poorly maps on complex multimorbidity [2]. However, in this excellent article “surgery” is even not mentioned.

RCTs must be properly designed to address a clearly defined issue (question) [3]. They should be randomized to avoid bias of inclusion and be double blind to avoid bias of evaluation by the patient and/or the physician. These criteria unfortunately can be difficult to meet. Randomization rarely is a problem, but blinding can be difficult or impossible. Although blinding is essential for all endpoints with a high placebo effect as in pain studies, medical treatment of endometriosis that affects menstruation is readily recognized by the patient as discussed previously [4]. A comparison between laparoscopy and laparotomy is readily recognized by the patient unless sham incisions are made, which would be ethically unacceptable. A major unsolved

The authors declare that they have no conflict of interest.

problem is how to interpret and use trials that are not perfect. This frequently becomes a concern in meta-analyses, when less than perfect RCTs are included, using the argument that they are the best available evidence. An example was the nonblinded “Endocan” study for surgery of superficial endometriosis [5] and the subsequent Cochrane analysis [6] that was withdrawn for that same reason [7].

An RCT, in which only 1 variable differs between the groups, is widely believed to be appropriate to evaluate efficacy and side effects. However, if a large trial is needed to establish efficacy, the clinical importance of the outcome risks is less important. The RCT balances between 2 extremes: Either the available clinical evidence is already so strong that it becomes almost unethical to perform the trial, or the expected benefit is so little that the result will be clinically of little importance. In addition, an RCT is not suited to evaluate rare events such as complications of surgery: To have 30 “complication cases” in each arm, an RCT of a 1% event would require a series of 6000 patients, which is not very realistic. In any case, an RCT is a slow method to establish evidence.

Another problem of RCTs is the inclusion and exclusion criteria that limit the applicability of conclusions to the group of patients matching these same criteria. Extrapolation of results to other age groups or those who in addition have specific diseases such as diabetes or hypertension becomes questionable. On the other hand, a trial without restrictive inclusion and exclusion criteria would require unrealistically high numbers of subjects to provide adequate numbers to randomize the smaller subgroups involved. Applied to surgery, and especially complex surgery, the heterogeneity of the population involved becomes a major problem to randomize a reasonably homogeneous number of patients. In addition, the individual preferences, techniques, and individual surgical skills that are variable limit multisurgeon trials. A study designed to compare different approaches or techniques, for example laparotomy versus laparoscopic approach, in addition to the variables cited, would have to consider the different skills of the surgeons involved on the 2 surgical approaches. Thus, the trial risks to evaluate the surgeon and his or her environment rather than the technique itself, “the singer or the song.” The mean duration of 150 minutes to perform a laparoscopic hysterectomy in the EVALUATE trial is an example of this [8]. These limitations of the RCT generated tensions with observational medicine [9,10].

Communication and exchange of ideas have changed profoundly over the last decades. The number of publications, the open access, and other sources of information have increased to a level that quality control of written information has become an issue [11]. Simultaneously, the exchange of expertise at meetings, live surgery, and individual discussions permits the individual opinion of a surgeon based on his or her expertise to enter the pool of a consensus opinion of the worldwide community of (expert) surgeons. This world community of surgeons is also sharply aware when

there is a lack of consensus. Any attempt to assess some of the variables in a rapidly evolving surgical field with an RCT risks to be outdated by the improved understanding and clinical evidence, even before it is published. In fact, a minimal period of 2 years is needed to design, perform, and publish the results of a trial. This is indeed slow when compared with the rapidity of the evolution of consensus opinions.

Let us take “surgical treatment of deep endometriosis” as an example. It is clearly evident that the knowledge on the disease itself, concepts of treatment, and surgical techniques have evolved continuously over the last 25 years and reached a consensus over many aspects of the condition, whereas others remain in debating stage. Given the absence of RCTs and almost nonexistent so-called hard evidence, the cumulative knowledge and agreement of a group of experienced surgeons on this condition had to be used to outline its definition, mode of diagnosis, and appropriate treatment [12]. As stated above, this was necessary for such a multifaceted difficult disease. Similarly, this consensus opinion was used recently to express concern that the severity and prevalence of deep endometriosis may be increasing [13].

The evidence value of this ubiquitous consensus opinion, as opposed to the individual surgeon’s opinion, is not yet reflected in the evidence ranking of traditional EBM. This is not surprising because this global consensus opinion has only become apparent recently with the impact of internet, social media, traveling, and interaction. Today, most surgeons know each other personally, as opposed to knowing each other from the literature, as was common in the past. Although poorly defined in the absence of rules or leadership, this consensus opinion has several advantages over RCTs. It is close to clinical applicability and it reflects the cumulative available experience of many surgeons, which also includes rare events such as complications in the entire population, even those with many complex morbidities. The speed of incorporation of new data precedes those of any trial by at least several years.

In medicine and surgery progress is achieved by the incorporation of knowledge and evidence obtained through research and clinical observation and experience that is shared, discussed, and approved at appropriate forums. That only some aspects have been demonstrated in randomized trials is not surprising given the ethical barriers to sham surgery, although this may be viewed by some as a problem. Although the scientific and statistical solidity of the well-performed RCT is beyond doubt, we must be fully aware of its limitations in surgery, especially in complex surgery. RCTs are not suited for rare events. Extrapolation of results must be done carefully, blinding is frequently impossible, and the numbers required for meaningful conclusions are often prohibitive. In meta-analysis the selection criteria to accept or reject nonperfect RCTs are another source of bias.

Therefore, observational medicine, clinical experience, and expert opinion continue to retain validity, and the conclusions should be maintained until proven otherwise with an

RCT, if applicable. Nonrandomized surgical series, documenting the surgical intervention, the outcome, and complications, remain valuable provided potential biases can be evaluated. Because retrospective series risk a recall bias, prospective series are superior. Besides the obvious surgeon bias, the methods used to prevent bias in the evaluation of outcomes are similar and as important as in RCTs. Blinding is important, especially for endpoints with a high placebo effect such as pain. In addition, the interpretation of many surgical outcomes of both RCTs and prospective series should take into account the impact of softer variables as patient psychology, personality, and attitude. Unfortunately, this type of information is rarely available. The long-term outcomes and changes of pain symptoms over time are also rarely available because most studies do not have a sufficiently long follow-up. However, the personal clinician, treating and following the patient over time, may have this information but not as part of a study. It is obvious that the preceding and the non-reported rare accidents and complications are the source of the so-called clinical impression of the surgeon. Therefore, observational medicine remains important.

Observational medicine recently reached a new dimension through modern communication, speed of information exchange, ease of travel, live surgery, and email discussions and interaction. Experience and the opinion of the individual has become, for many aspects of the question, a consensus opinion of many if not all (expert) surgeons worldwide. Moreover, the value of such a consensus opinion comprises clinical applicability for the entire population with all its complexity and is based on the widest available experience, including complications and mistakes. Consensus opinion therefore should be given a much more important ranking than “ideas and opinions” in the pyramid of evidence of EBM, and this should be reflected in our guidelines. The awareness of the clinical importance and evidence value of this rapidly evolving consensus opinion, especially for complex surgery, emphasizes the need to properly organize the collection of this knowledge and evidence within an international center or society.

Philippe R. Koninckx, MD, PhD*
 Department of Obstetrics and Gynecology
 Catholic University Leuven
 University Hospital
 Leuven, Belgium

Anastasia Ussia, MD
 Gruppo Italo Belga
 Villa del Rosario and Gemelli Hospitals Università
 Cattolica
 Rome, Italy

Errico Zupi, MD, PhD
 Department of Biomedicine and Prevention
 University of Tor Vergata
 Rome, Italy

Victor Gomel, MD
 Department of Obstetrics and Gynecology
 University of British Columbia
 Women's Hospital
 Vancouver, British Columbia
 Canada

*Corresponding author: Prof. Dr. em. Philippe R. Koninckx, MD, PhD, Catholic University of Leuven, University Hospital Gasthuisberg Dept of Obstetrics and Gynaecology Herestraat 49, 3000 Leuven, Belgium
 E-mail: pkoninckx@gmail.com

References

- Masic I, Miokovic M, Muhamedagic B. Evidence based medicine—new approaches and challenges. *Acta Inform Med.* 2008;16:219–225.
- Greenhalgh T, Howick J, Maskrey N. Evidence based medicine: a movement in crisis? *BMJ.* 2014;348:g3725.
- Jones DS, Podolsky SH. The history and fate of the gold standard. *Lancet.* 2015;385:1502–1503.
- Koninckx PR, Craessaerts M, Timmerman D, Cornillie F, Kennedy S. Anti-TNF-alpha treatment for deep endometriosis-associated pain: a randomized placebo-controlled trial. *Hum Reprod.* 2008;23:2017–2023.
- Berube S, Marcoux S, Langevin M, Maheux R. Fecundity of infertile women with minimal or mild endometriosis and women with unexplained infertility. The Canadian Collaborative Group on Endometriosis. *Fertil Steril.* 1998;69:1034–1041.
- Jacobson TZ, Barlow DH, Garry R, Koninckx P. Laparoscopic surgery for pelvic pain associated with endometriosis. *Cochrane Database Syst Rev.* 2001;CD001300.
- Jacobson TZ, Duffy JM, Barlow DH, Farquhar C, Koninckx PR, Olive D. WITHDRAWN: laparoscopic surgery for subfertility associated with endometriosis. *Cochrane Database Syst Rev.* 2014;CD001398.
- Garry R, Fountain J, Brown J, et al. EVALUATE hysterectomy trial: a multicentre randomised trial comparing abdominal, vaginal and laparoscopic methods of hysterectomy. *Health Technol Assess.* 2004;8:1–154.
- Redwine D, Mann CH, Wright JT. Evidence on endometriosis. Elitism about randomised controlled trials is inappropriate. *BMJ.* 2000;321:1077–1078.
- Farquhar CM. Extracts from the “clinical evidence”. Endometriosis. *BMJ.* 2000;320:1449–1452.
- Koninckx PR, Batt RE, Hummelshoj L, McVeigh E, Ussia A, Yeh J. The elephant in the room: quality control of endometriosis data. *J Minim Invasive Gynecol.* 2010;17:637–640.
- Koninckx PR, Ussia A, Adamyan L, Wattiez A, Donnez J. Deep endometriosis: definition, diagnosis, and treatment. *Fertil Steril.* 2012;98:564–571.
- Koninckx PR, Ussia A, Keckstein J, Wattiez A, Adamyan L. Epidemiology of subtle, typical, cystic, and deep endometriosis: a systematic review. *Gynecol Surg.* 2016;13:457–467.