In the absence of an adequate animal model, endometriosis remains a poorly understood disease. Its pathophysiology and epidemiology are topics of debate, the natural history is unknown, and treatment remains controversial [1]. “Endometrial glands and stroma located outside the uterus” as definition of endometriosis might not always represent a pathological condition [2]. Thus, another definition is needed for the disease causing pain and/or infertility and other symptoms. The prevalence and epidemiology are unclear because laparoscopy is necessary to make the diagnosis and because of the variable inclusion of subtle lesions [3]. Although Sampson’s theory remains the most widely accepted theory explaining the pathophysiology, evidence obtained over the last decade allows us to postulate that genetic or epigenetic changes are required for symptomatic endometriotic disease to develop, while the original cells are less important. These can be adult or neonatal endometrium, stem cells, or bone marrow–derived cells [4]. The type of genetic or epigenetic changes will determine the progression to typical, cystic ovarian, deep pelvic, or extrapelvic lesions; these can be considered as different diseases. Unfortunately, all types of endometriosis continue to be perceived as a single disease entity, owing to the concept of retrograde menstruation with implantation and progression of these normal endometrial cells due to their presence in an abnormal environment.

Medical and surgical treatment of endometriosis is hampered by the poor correlation between the severity of symptoms and the severity of lesions. Most trials have been performed with only a single type of therapy. Moreover, at meetings, alternative therapies are generally presented by 2 different speakers, often organized as a debate. The strengths of the evidence of each therapy are judged by statistical evidence and by the quality of the study design, and the conclusions are summarized in evidence-based medicine (EBM) guidelines. We should realize that this conglomerate of many little pieces of evidence suits our Western world’s Cartesian thinking that “understanding improves by dissecting a problem in small parts.”

Clinical medicine differs from research and requires a holistic approach, which unfortunately is practically impossible to realize in a huge comprehensive trial. Research on the pathophysiology, diagnosis, and treatment of endometriosis provides data, which are obviously useful. Clinical medicine is not limited to the results of specific trials, but encompasses all women and pathologies, including those with multiple morbidities. The necessary translation of research data into clinical medicine requires clear awareness of the limitations of evidence produced by trials. It requires an understanding of statistical significance and of the absence of evidence together with an awareness of the limitations of each trial design, which balances between ideal and realistic, owing to the human bias confronted with massive information. We do indeed judge the available data, based on our educational background during the (limited) time available to us, and most likely colored by our beliefs. The conscious awareness of how each of these elements are influencing our judgment might provide valuable insight into the medical and surgical treatment of endometriosis, which has recently been characterized as a “100 Years War” [1].

Materials and Methods

A systematic review of each of the more than 40,000 articles in the literature is realistically impossible, and the result would be clinically irrelevant because of the numerous disparate opinions. To restrict the number of articles according to some predefined elements of quality, as is done in most reviews and meta-analysis or guidelines, is not a solution, because occasional accidents, complications, and case reports will be missed. Consequently, we drafted this paper as a critical appraisal of our clinical practice in treating women with endometriosis as developed by reading the literature, by attending congresses, and mainly after many hours of discussion.

Statistical Evidence and the Population

A law of physics is evidence, based on repeatable observations without a single exception in the given circumstances. The oft-heard statement that there is no “evidence” for a parachute or a guillotine does not take into account that this evidence belongs to this type of law-evidence.

Statistical evidence of an effect is based on the probability that an observed difference is true. This statistical evidence,
known as “significance,” remains a probability, comprising spurious significances and a (small) probability of error. Significance increases with the number of observations; therefore, absence of statistical difference does not permit a conclusion. Conclusions are limited to the group of subjects investigated, that is, matching exactly the inclusion and exclusion criteria, such as age, blood pressure, weight, etc. Extrapolating results to the entire population is always hazardous and based on judgment. To judge the effect of a therapy, a statistically or mathematically significant difference, although an obvious prerequisite, is not sufficient for clinical usefulness, because this requires some magnitude of the effect.

The clinical usefulness of diagnostic tests is subtler. Their value is assessed statistically by their sensitivity and specificity for a given population, whereas their clinical usefulness of these values must be judged clinically. Clinical judgment indeed is needed to determine whether a test with 10% false positives and 10% false negatives (i.e. with 90% sensitivity and 90% specificity) is clinically useful. Statistical differences should not be inverted as useful predictors; for instance, although men are significantly taller than women, height is a poor predictor of sex. Clinical usefulness of a test often needs clinical stratification of prediction. Indeed, the statement that “ultrasound highly accurately predicts the presence of a deep endometriosis nodule” has limited clinical usefulness. Although true for all nodules taken together, this is not necessarily true for all sizes of nodules and does not take into account the lower detection limit. A third limitation is that any investigated population may hide a smaller subpopulation with a different or opposite effect. An example from deep endometriosis is that it took 20 years before we realized that a small subgroup experienced bowel perforations during pregnancy [5], in contrast to the regression expected during pregnancy.

The limitations of evidence, clinical usefulness, and statistical significance hold true for all forms of statistics as well. The many pitfalls of statistics include the subtle and likely unintended misuse of statistical analysis [6–8]. In conclusion, statistical evidence of differences is very useful in research, but does not necessarily permit an assessment of clinical importance, which is based mainly on the magnitude of the (significant) effect and on the specificity and sensitivity of a specific diagnostic question.

Prerequisites for statistical evidence include identical populations and the absence of bias in observation. Population bias requires strict randomization. For easily measurable endpoints, such as height, weight, and pregnancy, observation bias is minimal. For other types of endpoints, such as pain and well-being, observation biases can be strong. These are known as placebo effects and observer bias. Therefore, double-blind trials are mandatory to obtain evidence for these endpoints. Statistical analysis cannot make up for poor data collection; the level of a river cannot rise above that of its source.

The EBM pyramid of evidence is fully mathematically valid, with the confirmed double-blind randomized controlled trial (RCT) on top of the pyramid. Unfortunately, however, the many restrictions are rarely highlighted. A first obvious restriction is that a significance of 0.05 means a 95% probability of being true but also a 5% probability of being false. A second restriction is that the most perfect randomization of the population cannot ascertain a homogeneous effect since a small hidden population with a different/opposite effect will not be detected. A third major problem is that huge trials are necessary to detect rare events. Indeed, an event occurring in 1% requires a prospective trial of roughly 3000 women to find 30 events, the number needed for meaningful statistics. The fourth and most important problem is that so many important variables are involved in endometriosis and its therapy that a trial taking all of them into consideration would be unrealistic. This problem is similar to that of multiple morbidity, which is not suited for a RCT.

The Pandora’s box of EBM and endometriosis treatment comprises all forms of incorrect use of evidence [6–8] and of ignoring the limitations of trials. Accidents, such as someone tripping on cables in the operating room, causing an electrical blackout with injury to the patient, are rare events with solid evidence of a causal relationship between cause and effect; however, this will not be picked up in an RCT. Randomization to achieve comparable populations and double blinding for pain or well-being trials are important for medical and surgical therapy. The value of a nonblinded trial for pain or well-being as endpoints is questionable. Rare events and complications or accidents will be picked up only by observational medicine. Moreover, surgical trials inherently involve important additional variables associated with the surgeon and quality of surgery.

The “Player” and the Endometriosis Bias

As a consequence of increasing subspecialization and the sheer amount of data available today, the all-round gynecologist with mastery of both advanced surgery and basic endocrinology no longer exists. Even those of us trained initially in one aspect of the discipline will rapidly become less up to date after moving to another one. Like a mother loves her children, we all have the biases of our interests.

The unspoken reality is that each of the many factors, including environment, ambition, work, hazards, publications, grants, congress presentations, matter for our careers. Affiliated with university or industry, young or old, researchers and academics alike have common interests in patents, publications, and presentations. This translates into research grants, travel grants, visibility, and publications that ultimately support promotions and/or private practice.

Endometriosis has 2 specific biases. The severity of disease often becomes fully apparent only during surgery. The absence of a validated classification hampers comparisons of results. The Pandora’s box of endometriosis includes the margin of error when the diagnosis is made without resorting to laparoscopy and the absence of a validated classification system. The player’s bias contains all forms of imperfect data and
their interpretation due to personal interests. Given that data manipulation can be insidious and that the “honest” removal of outliers can be subjective, intention-to-treat analysis was introduced. Both surgical and drug companies assist individuals and societies with research and travel grants, and provide support to congresses. The budgets of surgical companies are relatively small, and these companies are rarely directly involved in trials. In contrast, all major trials of medical treatment of endometriosis reported to date were organized by the pharmaceutical industry. Although these trials were scrupulously randomized and monitored, the subtle manipulation of trial design is rarely discussed; examples include the choice of comparator drug and noninferior analysis. A much greater problem is the fact that many trials with unfavorable results are either stopped after interim analysis or not published, as evidenced by an analysis of registered RCTs [9]. This difference in support provided by the surgical industry is a reason why surgical trials are smaller and often poorly monitored. Fortunately, the overall integrity of both medical and surgical gynecologists is high.

Medical Therapy Revisited

Medical therapy for endometriosis is not useful for infertility [10]. To the best of our knowledge, the Anti-TNFa trial [11] is the only double-blind RCT to date evaluating the pain associated with endometriosis. Remarkably, in that trial there was a very strong placebo effect; some women who previously needed monthly morphine injections became almost pain-free. Unfortunately, all trials affecting menstruation were not blinded, since the women were aware of their menstruation. Another specific bias of all major endometriosis trials is the inclusion criteria, which used to be pain and laparoscopic and/or histologically proven endometriosis in the last 1, 2, or 3 years. However, during the diagnostic laparoscopy necessary to confirm endometriosis, in most cases all visible superficial and cystic lesions of endometriosis are excised or coagulated. In such cases, whether the remaining pain is still caused by endometriosis would be questionable.

Nonetheless, this question does not completely invalidate the numerous reports that medical treatment decreases endometriosis-associated pain [12–14]. A decrease in pain indeed seems logical, considering that the endometrium stops growing and decidualizes in the absence of estrogens and/or presence of progestogens. Less well documented are the long-term effects [15] in terms of prevention of progression or recurrence. Data are limited to a slightly lower incidence of typical lesions after years of treatment with oral contraceptives and lower recurrence rates of cystic ovarian endometriosis during treatment. There are no data indicating that medical treatment prevents the onset of endometriosis or the progression of subtle lesions, deep endometriosis, or extragenital endometriosis. There also are no data suggesting that medical treatment prevents progression in all women, with some cases of severe endometriosis persisting even after more than 10 years of medical treatment (personal observation). It is surprising that no attention has been given to the effect of medical therapy on the steroid hormone concentrations in peritoneal fluid, a space in which peritoneal endometriosis grows.

The Pandora’s box of medical therapy includes that fact that in nonblinded medical trials with gonadotropin-releasing hormone, progestins, or estroprogestins, the placebo effect has been poorly investigated if not ignored. Moreover, observer bias has not been investigated. The inclusion criteria of most trials call into question whether the pain that these women were experiencing was due to endometriosis after undergoing surgery. This contrasts sharply with the widely held belief in the efficacy of the recommended life-long treatments to treat pain and prevent progression. Another major problem is the frequent use of medical therapy for longer periods in women with pain and suspected endometriosis, but without a documented diagnosis.

Surgical Therapy Revisited: Individual Experience

Laparoscopy is necessary to diagnose superficial endometriosis, during which typical lesions can be excised, vaporized, or coagulated [16]. The proof of efficacy of this treatment is limited to one double-blind RCT for pain [17]. That trial found a significant placebo effect persisting for several months, but with a widely variable magnitude of effect. The effectiveness of surgical treatment of superficial endometriosis to increase infertility is unclear. Indeed, the Gruppo Italiano did not find an increase in fertility. In contrast, in the nonblinded Endocan trial, whether pregnancy rates increased after surgery, or whether pregnancy rates decreased in the control group as a result of the stress caused by the awareness that the endometriotic lesions had not been removed, was not clear [18]. Nevertheless, in view of the low surgical risk associated with the procedure, not to treat is not an option.

Cystic ovarian endometriosis is diagnosed by ultrasound. After surgery (excision or vaporization), spontaneous cumulative pregnancy rates are 50% to 60% and recurrence rates range from 5% to 20%. Undertaking a trial designed to evaluate whether surgical treatment and adhesiolysis affect progression would be ethically questionable. The outcome of a surgical intervention seems to vary with the surgeon [19]. A special problem is posed by small cystic ovarian endometriomas, especially in young girls [20]. Indeed, because both endometrioma and the surgery to remove it can damage oocyte reserve, and considering a recurrence rate of 5% to 20%, the decision to perform surgery must strike a balance between the risk of an enlarging cyst and the risk of repeat surgery. Although poorly defined, deep endometriosis is associated with severe pain in most women and severe bowel and ureter problems in some. The natural history is not known, but clinical observation suggests that most lesions are no longer progressive at the time of diagnosis. However, some such lesions may progress rapidly. The added value of ultrasound and MRI for the diagnosis and the radicality of surgical
excision or the need of bowel resection remains debated. Notwithstanding the reported postoperative spontaneous cumulative pregnancy rate of 20% to 50%, whether surgery improves fertility is unclear. Pain relief is well documented, albeit not in RCTs [21], the feasibility and ethical aspects of which would obviously be questionable. The reported rate of recurrence of deep endometriosis nodules is <1%, whereas recurrence of pain and of subsequent surgery is at least 20%. Deep endometriosis surgery is difficult and prone to complications.

Besides the technical aspects of the procedure, the surgeon’s experience with and knowledge of endometriosis are important. The patient’s symptoms, preoperative imaging findings, and expectations as revealed by preoperative counseling modulate the choice of intervention performed. Although all surgeons are strongly aware of this aspect, it cannot be found in publications. Moreover, this aspect would not be compatible with an RCT.

Another element in the Pandora’s box of surgical treatment of endometriosis is the absence of quality control. Indeed, without video registration, the diagnosis, the completeness of excision (especially from the bowel or ureter), the ovarian damage caused, and the skill of the surgeon cannot be judged [22]. The latter has become even more important, given that the duration of surgery and the extent of manipulation have been identified as key factors in adhesion formation [23]. The fact that preoperative findings will influence choices made during surgery is an element to consider in multidisciplinary approaches.

Sequential Therapy

Surgical therapy can be followed by medical therapy. Although this approach is widely used to prevent disease progression, data on its effectiveness are limited. Moreover, any data gathered are hampered by the absence of information regarding the completeness of surgical diagnosis and surgical treatment without video registration.

Medical therapy can be followed by surgery. Clinical observation of very severe endometriosis after many years of medical therapy in women with progressive symptoms suggest that endometriosis has been progressive in these women. Moreover, a frozen pelvis is a common observation after repetitive in vitro fertilization (IVF) cycles in women with deep endometriosis. In young symptomatic women, how to balance early diagnosis and surgery vs medical treatment is often unclear. In the absence of data, this discussion seems to be based on whether endometriosis is considered a recurrent disease after complete excision. If not, then early surgery seems preferable; otherwise, it seems wise to postpone surgery for as long as possible.

The Clinical Reality of Endometriosis Treatment

The clinical reality of endometriosis begins with the frequent delay in diagnosis, either because endometriosis is not considered when clinical examination and ultrasound findings are negative or because of a reluctance to perform diagnostic laparoscopy, especially in young women [24]. This reluctance may be fueled by the perception that the quality of surgery is variable, that surgery can cause ovarian damage, that the skills to treat unsuspected severe endometriosis are not always present, and that the quality of surgery is difficult to evaluate.

Surgery for superficial pelvic endometriosis is considered mainstream. Surgery for cystic ovarian endometriosis is technically difficult [25], even though ovarian cysts were once erroneously considered the first level of surgery by such bodies as the Royal College of Obstetricians and Gynaecologists. The technical difficulty of larger deep endometriotic lesions is spurring a shift toward pelvic surgeons and technically skilled oncologists and abdominal surgeons with limited knowledge of the disease.

Medical therapy is widely used in women with suspected endometriosis to avoid surgery and/or to prevent progression. It is widely used after surgery to prevent recurrence or because of incomplete surgical excision. The increased success rate of IVF has led to frequent use of IVF even before diagnostic laparoscopy is performed. The end result is that the information provided to the patient and the therapy given vary widely based on the surgeon’s background. The cost of treatment is only now beginning to be addressed [22].

The Pandora’s box of clinical reality includes the fact that endometriosis specialist care is spread over fertility specialists, medical treatment specialists, and surgeons. Referrals occur, but are less frequent than needed. A major problem is that the available evidence does not permit clear conclusions. All the 3 types of specialists noted above are undoubtedly honest, notwithstanding their differing opinions and approaches to treatment.

Conclusions

Statistical evidence, including the RCT, is the probability that the means of 2 populations or therapies are different. Since the accuracy of the estimation of the mean increases with sample size, significance increases with sample size. Mathematical significance however may not indicate clinical usefulness since this requires some magnitude of the effect. In addition, results are valid only for the group investigated as defined in inclusion and exclusion criteria. The published data on the outcomes of endometriosis treatment are not very clear, unfortunately. The view of medical treatment versus surgical treatment of endometriosis as a “100-year war” stems from the fact that the weaknesses and limitations of each type of therapy are rarely addressed in the title and abstract of original articles but are often buried in the discussion.

The limitations of the noninvasive diagnosis of minor forms of endometriosis remain major problems. Symptoms do not always reflect disease severity, and biochemical markers and imaging still are not very useful. For cystic ovarian and deep
endometriosis, a combination of complete anamnesis, ultrasound investigation with high-resolution machines, and magnetic resonance imaging by well-trained radiologists may provide information to guide the choice of further therapeutic steps. Solid evidence on the treatment of pain associated with endometriosis is limited. Indeed double blind trials are needed for endpoints as pain and well-being because of the placebo effect and the observer bias. This unfortunately is not possible for medical therapies when menstruation is affected or for surgery for ethical reasons. For endometriosis therapy, publication bias is huge and quality of information is low. For medical therapy the main problems are the non-published trials and the huge commercial interest in trial outcome. For surgery the main problem is the limited numbers of interventions by each surgeon, the variability in surgical techniques, in surgical skills and in complexity. A single center trial therefore risks not having sufficient power whereas a multicenter trial risks evaluating rather the surgeon than the intervention. Information on accidents and rare events and are limited to case reports and observations. All various surgical techniques should be available in the same institution to avoid a bias due to a lack of surgeons with proper training and expertise. An option could be to create centers of excellence with expertise in every aspect of the disease: diagnosis, medical and surgical treatments, and proper patient follow-up.

It would be useful if we could focus on what we agree on, without being polemic. This could serve as the basis of the information given to women with pelvic pain, infertility, or endometriosis. Based on the available evidence, we can draw the following conclusions: (1) a woman with pain and/or infertility has a 50% probability [26] of having typical endometriosis or worse; (2) prolonged medical treatment of endometriosis without a diagnosis is not recommended; (3) superficial endometriosis can be diagnosed only by laparoscopy; (4) medical therapy for endometriosis can reduce pain but is ineffective for infertility and for cystic ovarian endometriosis; (5) whether medical therapy prevents progression of deep endometriosis in all women is unclear; (6) diagnostic laparoscopy should be recorded to permit subsequent confirmation of diagnosis and of completeness of diagnosis; (7) the possibility of treating the disease during diagnostic laparoscopy is a plus; (8) quality control of surgery is only possible with video registration of the entire intervention; (9) informed consent requires that the patient be given correct information on the indication, planned intervention, and surgeon’s experience; (10) EBM should be based on the best evidence available, including rare events and the limitations of RCTs; and (11) our actual clinical management approach based on experience should be maintained unless proven otherwise.

References


Philippe R. Koninckx, MD, PhD
Anastasia Ussia, MD
Jorg Keckstein, MD, PhD
Leila Adamyan, MD, PhD
Errico Zupi, MD, PhD
Arnaud Wattiez, MD, PhD
Victor Gomel, MD

aDepartment of Obstetrics and Gynecology, Catholic University Leuven, University Hospital, Gasthuisberg, Leuven, Belgium
bGruppo Italo Belg, Villa del Rosario and Gemelli Hospitals, Università Cattolica, Rome, Italy
cEndometriosis Center, Villach, Austria
dUniversity of Ulm, Ulm, Germany
eUKT, University Hospital of Tübingen, Tübingen, Germany
fV. I. Kulakov Research Center for Obstetrics, Gynecology, and Perinatology, Ministry of Health of the Russian Federation, Moscow, Russia
gDepartment of Reproductive Medicine and Surgery, Moscow State University of Medicine and Dentistry, Moscow, Russia
hDepartment of Biomedicine and Prevention, University of Rome Tor Vergata, Rome, Italy
iDepartment of Obstetrics and Gynecology, University Hospital of Strasbourg, Strasbourg, France
jDepartment of Obstetrics and Gynecology, Latiffa Hospital, Dubai, United Arab Emirates
kDepartment of Obstetrics and Gynecology, University of British Columbia Women’s Hospital, Vancouver, BC, Canada

*Corresponding author: Philippe R. Koninckx, MD, PhD, Vuilenbos 2, B 3360, Bierbeek, Belgium. E-mail: koninckx@gmail.com


