Hierarchy of evidence for endometriosis diagnosis and surgery.

Philippe Koninckx[1], Anastasia Ussia[2], Joerg Keckstein[3], Leila adamyan[4], Jacques Donnez[5], Dan Martin[6], and Arnaud Wattiez[7]

[1]KU Leuven
[2]University Hospital Agostino Gemelli
[3]University of Ulm
[4]Moscow State University
[5]Universite catholique de Louvain
[6]University of Tennessee Health Science Center
[7]Universite de Strasbourg Faculte de Medecine

May 16, 2022

Abstract

Without an animal model and a non-invasive diagnosis, the pathophysiology of endometriosis is unclear and information is limited to symptomatic women. Lesions are biochemically variable. Medical therapy cannot be blinded and extensive surgery combines low numbers with variable difficulty and surgical skills. Experience is spread among specialists in imaging, medical therapy, infertility, pain and surgery. Besides the recent changes in interpreting statistical analyses, the limited good-quality evidence increases the importance of clinical experience. Therefore trial design, analysis and judgment of results should be done by experts in the different disciplines of endometriosis, before being translated into guidelines.

Hierarchy of evidence for endometriosis diagnosis and surgery.

Koninckx Philippe R.[1,2]

Ussia Anastasia[3]

Keckstein Jörg [4]

Adamyan Leila[5]

Donnez Jacques[6]

Martin Dan C. [7]

Wattiez Arnaud [1,8]

[1] Latifa Hospital, Dpt OBGYN, Dubai, Unite Arab Emirates

[2] Prof emeritus OBGYN KULeuven Belgium, University of Oxford, and Hon Consultant, UK, University Cattolica, Roma, Italy and Moscow State Univ, Russia.

[3] Consultant Università Cattolica, Dpt OBGYN Roma Italy

[4]Endometriosis Centre, County Hospital Villach, Austria and University Ulm, Dpt OBGYN, Ulm, Germany

[5] Department of Operative Gynecology, Federal State Budget Institution V. I. Kulakov Research Centre for Obstetrics, Gynecology, and Perinatology, Ministry of Health of the Russian Federation, Moscow, Russia; and Department of Reproductive Medicine and Surgery, Moscow State University of Medicine and Dentistry, Moscow, Russia

[6]prof emeritus Catholic University of Louvain, Dpt OBGYN, Brussels, Belgium

[7] Professor Emeritus, Dpt OBGYN, University of Tennessee Health Science Centre, Memphis Tennessee, USA; Institutional Review Board, Virginia Commonwealth University, Richmond, Virginia, USA

[8] Prof Department of obstetrics and gynaecology, University of Strassbourg, France

Corresponding author

Philippe R. Koninckx

Vuilenbos 2

3360 Bierbeek

Pkoninckx@gmail.com

Running title

Hierarchy of evidence in endometriosis.

Abstract

Without an animal model and a non-invasive diagnosis, the pathophysiology of endometriosis is unclear and information is limited to symptomatic women. Lesions are biochemically variable. Medical therapy cannot be blinded and extensive surgery combines low numbers with variable difficulty and surgical skills. Experience is spread among specialists in imaging, medical therapy, infertility, pain and surgery. Besides the recent changes in interpreting statistical analyses, the limited good-quality evidence increases the importance of clinical experience. Therefore trial design, analysis and judgment of results should be done by experts in the different disciplines of endometriosis, before being translated into guidelines.

Introduction

Medicine started with observations, experience, and trial and error, initially hampered by conflicting beliefs such as Greek mythology. Later, statistical analysis was needed to interpret and evaluate the variability of observations and to help understand mechanisms. For more than a century classical statistical methods used significance levels, power and P-values. However, the interpretation of the widely used P-values has created confusion (1). P-values were introduced 100 years ago by R Fisher to grasp in one value the probability that an observed effect, can be attributed to chance (null hypothesis), taking into account its distribution and power (2). A P-value thus measures the extremeness of a result given the null hypothesis but not that the hypothesis is true given the observed results. Judgement of the validity of a hypothesis requires Bayesian reasoning, needing a prior probability and the Bayesian factor indicating the improvement of that probability (3). Instead of being an argument to accept a hypothesis, a P value of 0.05 or 0.01 only increases a prior probability of 50% to 71% or 89%, the initial hypothesis being wrong in 29% or 11% respectively (1). That the probability of a hypothesis depends on much more than P-values (4) led to the suggestion that many research findings are wrong and often 'an accurate measure of the prevailing bias' (5). The precision of P-values (6), the poor reproducibility (7) and the misuse in medicine resulted in 2016 in a warning by the American statistical association that P-values 'do not measure the probability that the studied hypothesis is true or the probability that the data were produced by random chance alone and that they are not a good measure of evidence regarding a model or hypothesis' (8). Examples illustrating the difference between P-values and probability are: the high P-value that measles produces a rash, cannot be used to predict the probability of

measles in those with a rash; similarly, a 60% probability of rain is different from a non-significant P-value describing the odds of rain. However, today, statistical reporting is changing slowly for most gynecological journals (9) except JMIG and BJOG (10, 11).

Evidence-based medicine (EBM) intended to integrate research data corrected for biases into clinical medicine and was initiated when calculators permitted more complex analyses such as meta-analysis. In the 1990ies, EBM embraced P-values for evidence resulting in a pyramid of evidence (12) with the RCT (13), and later their meta-analysis (14) and systematic reviews on top. However, the integration of results in clinical medicine and guidelines proved to be difficult (15). Besides that it is unclear how clinical experience is used to translate results into grades of evidence (15), this difficulty could be seen as the unconscious conflict between the inappropriate use of P-values as 'evidence', and clinical medicine using a rather Bayesian approach.

An EBM approach to endometriosis needs specific considerations. Without an adequate animal model and the ongoing debate about pathophysiology, it remains unclear whether endometriosis is one or several diseases. Without an adequate non-invasive diagnosis, epidemiology is poorly understood. RCTs are difficult to organise when medical therapy cannot be blinded since recognised by the patient (e.g. when affecting menstruation) and extensive surgery is too variable for the available numbers.

Therefore some difficulties with EBM will be reviewed to discuss EBM of endometriosis from a clinical perspective.

Evidence Based Medicine.

'Evidence-based medicine is the conscientious, explicit, and judicious use of current best evidence in making decisions about the care of individual patients' (16). As reviewed elegantly (12) evidence needs to be obtained by credible processes taking into account the quality and the totality of the evidence. However, the resulting hierarchy of evidence was often poorly reproducible and many systematic reviews could be considered a review of evidence. Therefore grades of evidence were introduced to judge all available evidence including observational series and case reports.

EBM can claim many successes in changing clinical therapy by formalising RCTs, reporting, meta-analyses and systematic reviews, in the prevention of selection and observation biases to obtain accurate data (the water of a river cannot rise above its source (17)), and in a rigorous clinical interpretation. EBM, initiated in the 1990s, embraced traditional statistical analyses and P-values to judge the efficacy of treatments, and the accuracy of diagnostic tests and their clinical interpretation to formulate guidelines and recommendations. However, the integration of EBM into clinical medicine has been difficult (12). Without reviewing the risks of misuse of P-values and the more recent Bayesian (18) approach to judge the validity of hypotheses, we need to realise that Bayesian inference has not yet been fully incorporated in EBM, although more similar to clinical decision making for diagnosis and therapy.

Clinical considerations of traditional analysis of treatment and diagnosis

Rarely considered is that traditional statistical analyses require a homogeneous population and that the analysis is not suited to detect smaller subgroups with different behaviour (19) or to handle rare events, that would need prohibitively large groups. Although well known that small, clinically irrelevant differences can reach 'significance' since P-values improve with the square root of the number of observations, this still creates a publication bias that is difficult to deal with (20, 21).

Diagnostic tests are used to estimate the probability that a patient has or does not have a disease, which are the positive (PPV) or negative predictive values (NPV). However, the accuracy of prediction decreases sharply when the prevalence of the diseases is low, especially below 5 or 1%. A Bayesian approach (18) permits the exact calculation of this relationship (PPV=sensitivity*prevalence/(sensitivity*prevalence + (1-specificity)(1-prevalence)). Clinically we knew this since a test with 99% sensitivity and 99% specificity for a disease with a 1% prevalence, has a predictive value of only 50% since the numbers of false positives as true

positives are equal. Unfortunately, prevalences are not always well known and can be variable e.g. because of a referral bias. Therefore, predictive values of diagnostic tests for rare diseases are better in tertiary referral centres (22) with a higher prevalence. The influence of prevalences on PPV's is still poorly integrated into the clinical interpretation of diagnostic tests. More difficult is the estimation of the combined diagnostic accuracy of several tests (23). The difficulty was illustrated in a recent Cochrane review, suggesting to use tests sequentially, using first a test with high sensitivity, and subsequently using another test to re-test the negative group (24). The added value and the combined accuracy of tests can be calculated with a Bayesian approach as demonstrated for endometriosis (25, 26), but this is overall rarely used.

Clinical judgment, experience and artificial intelligence

Difficult to standardise is the experience-based clinical judgement when types of data cannot be compared. It is not clear how to balance data such as the efficacy of treatment with severity and incidence of side effects. Moreover, these are not always independent variables, such as when both are surgeon dependent. A historic example is that chloramphenicol, an excellent antibiotic, was no longer used after many years because of the 1/10.000 risk of aplastic anaemia (27). It is not clear how to judge drugs when the absence of blinding and placebo effects need to be balanced. Also, the judgement of imprecision, inconsistency, publication bias and external validity, remains difficult. The same holds for the grades of evidence (15, 28), defined as the expectation that the conclusions will or will not be modified by further research, and for meta-analyses, requiring inclusion and exclusion criteria judging the quality of RCTs (29) up to becoming misleading (30). The value of clinical judgement is also illustrated by the conclusion that evaluation of biases in diagnostic accuracy by Quadas tools (31, 32) is not much superior to clinical judgment.

Clinical experience integrates knowledge with experience in the entire population, including heredity, age, antecedents, rare events and multimorbidity. This also includes the many conclusions based on common sense or previous experience without trial evidence. Trials are not performed when results have no clinical consequences, or when practically difficult to perform because of low prevalence as in multimorbidity or when practically impossible to perform. An example is the choice of suture material, which is based on tensile strength and resorption rate but without trial evidence, since suture or knot complications are rare. Many individual and local preferences were implemented following rare events, accidents or near accidents. Since these events are often forgotten years later it is suggested to be prudent when changing habits because of lack of evidence.

Clinical experience is complex (33) and some aspects are difficult to quantify such as the skill of a surgeon and rare but severe complications not reflected in RCTs. A clinical diagnosis considers a series of potential diagnoses, ranging from likely to rare, taking into account age, antecedents, symptoms, clinical exams, blood tests and imaging. The integration of all results into a PPV for each diagnosis including the risk of mistakes, is a complex experience-based, artificial intelligence like (34) process. Clinical experience also precedes RCTs (Fig 1) either when performed to confirm observations, or not performed when the superiority of a treatment or an intervention seems repetitively observed, without exceptions, or when the expected effect is so little that the result will be clinically irrelevant.

Emotional intelligence is rarely considered since even more difficult to judge. The interaction of the clinician and the patient through body language and expectations influences diagnosis and therapy, and similar data can be interpreted differently by clinicians with comparable experience. Therefore the attitude of patients is reflected, although not explicitly, in clinical judgment.

Conclusion

These considerations explain the ongoing discussions on the hierarchy of evidence(35), procedural aspects such as financial bias in funding (36) and drug research (37), and the epistemological discussion to distinguish justified belief from opinions (38). Also, medicolegal aspects influence clinical judgment as evidenced by the recent introduction of NUTS (Number of Unnecessary Tests to avoid one Suit) statistics (39). Clinical

experience and judgment somehow integrate the entire population with rare cases and multimorbidity with case reports and the vast literature of descriptive observations, influencing our judgement.

EBM and Endometriosis

Endometriosis is poorly understood

Endometriosis is a frequent disease causing pain and infertility and is the most frequent reason for surgery in women (40). Given the likely association with adenomyosis and bleeding disorders (41), endometriosis can be considered for almost any complaint in gynaecology. Without an animal model permitting experiments and without a non-invasive diagnosis, the pathophysiology, the natural history and the epidemiology (42) are poorly understood, and the data on endometriosis is scanty in adolescence and limited to symptomatic women. The consequent difficulty to handle the epidemiology of endometriosis is illustrated by the recent suggestion to redefine endometriosis as 'symptomatic', eliminating those who did not undergo a laparoscopy (43). Even for cystic ovarian endometriosis the accuracy of imaging seldom exceeds 90% while it is difficult to exclude ovarian cancer, especially in older women (44).

Good-quality data are limited

A laparoscopy is performed only in women with pain or infertility, and it is difficult to ascertain in individual women that endometriosis and pain or infertility are causally related since only half of the superficial lesions are painful (45) and considering the many other causes of pelvic pain or infertility. Medical therapy has an important placebo effect (46), but cannot be blinded since the patient recognizes active therapy, especially when affecting menstruation. It is unclear if women with 'proven endometriosis' as a criterion to be included in trials, still have endometriosis since they underwent laparoscopy with surgical destruction of endometriosis in most women. This variable judgment of a trial is illustrated by the not-blinded ENDOCAN trial (47) showing improvement in fertility following surgery, resulting in a Cochrane meta-analysis which was subsequently withdrawn (48). In cystic ovarian endometriosis the results of surgery, ovarian damage and recurrence rates are surgeon dependent (49). Deep endometriosis is highly variable and surgery is technically difficult and complication prone. Because of the variable skill of the surgeon, with low numbers of interventions, RCTs are not realistic, and if performed nevertheless, unexpected results risk being criticised as occurred recently in the LACC trial demonstrating higher recurrence rates after laparoscopic surgery for cervical cancer (50).

Clinical judgment varies with subspecialties

The clinical judgment of the evidence is further complicated by different sub-specialists. Clinical experiences are bound to vary between radiologists performing MRI, gynaecologists specialising in ultrasonography or endocrinology or medical therapy or surgery, and abdominal surgeons with little expertise in other aspects of endometriosis management. An additional difficulty is the degree of commercialisation and industrialisation (51), especially in Infertility and medical therapy.

Clinical judgment varies with our perception of pathophysiology

Management should be based on evidence, but the clinical judgement might vary with understanding pathophysiology. The implantation theory (52, 53) defined endometriosis as 'endometrial glands and stroma outside the uterus' and thus as one disease, which became clinically considered progressive and recurrent. According to the genetic-epigenetic (G-E), theory endometriosis starts developing after a cumulative series of cellular incidents (54). Endometriosis lesions thus are clonal and individually different which is consistent with the variable aromatase activity and progesterone resistance (55), and the variable response to medical therapy (56, 57). If lesions are different, traditional statistical analysis is inadequate (19). The risk of G-E incidents increases by oxidative stress of retrograde menstruation or the peritoneal microbiome. Therefore susceptible women have an increased risk after puberty, and the remaining group will have a progressively lower risk (58). Age thus becomes an important factor in epidemiology. Pelvic endometriosis lesions grow

in the peritoneal cavity which is endocrinologically and immunologically a specific microenvironment, but the growth of endometrial lesions is self-limiting (58) probably as a consequence of fibrosis and inflammation secondary to the immunologic reaction. This is consistent with the clinical observation that most deep endometriosis lesions that are followed clinically since symptoms were insufficient for surgery, do not grow. Viewed as a G-E driven disease, recurrences might become preventable by decreasing oxidative stress when decreasing retrograde menstruation. This is consistent with the lower recurrence rate of cystic ovarian endometriosis when taking oral contraception. Although not demonstrated yet, we might consider prevention by preventing ascending infections, or by changing the peritoneal microbiome by food intake and exercise (59). This is consistent with the observations that the risk of developing endometriosis seems lower when taking food rich in antioxidant as omega3, Vit E, Vit C and citrus (60, 61). It is too early to fully understand the effect of vitamins on inflammation and immune response in endometriosis (62). New concepts of pathophysiology should be considered for future trials. This could apply more specifically to endometriosis in adolescence, to the prevention of endometriosis, and to interpret results of endometriosis if heterogeneous and more than one disease.

Non-biomedical health systems

A growing number of reports document the management of endometriosis with complementary therapies (63-65), acupuncture, food intake(66) and exercise (59), and more recently traditional Chinese medicine(67). These reports are difficult to interpret since indications and results of treatment poorly fit EBM standards. However, indirect and circumstantial evidence is too strong to be ignored altogether.

We are at the crossroads of understanding the role of food intake and exercise on the peritoneal (68) and the intestinal microbiome. Both might influence endometriosis onset and growth either directly or through immunology and oxidative stress.

Conclusion

In conclusion, (Fig 1) high-quality evidence is limited and the clinical judgment varies with experience, which is different for each subdiscipline involved. Without questioning the importance of the rules to grade evidence, recognise bias and understand statistical analysis, either numerical or Bayesian, the differences in judgment by subdisciplines need to be addressed. It seems logical that the ranking of evidence for diagnosis, medical therapy and surgery should be performed separately by different subspecialists. Although the lines between disciplines are not rigid, the prior hypotheses to be tested must be formulated by clinicians with experience.

Surgery for severe deep endometriosis needs specific comments. Data are limited to observational series with referral biases and differences in technique. However, the surgeons with extensive experience (e.g. more than 5-10 years of experience and more than 200 interventions) are a small group, who know each other's surgery and who meet and discuss several times a year and progressively adapt their surgery (69). Therefore the elements on which this group agrees because of a similar experience not changing over time, constitute rather solid evidence. Hopefully, statisticians will help to formalise these experience-based observations of this group into evidence. Similarly, it seems important to register when and why opinions/experiences are different.

Discussion

The principles of EBM (16) are clear, but the ranking of the evidence is struggling with a poorly defined clinical judgment and experience, which for endometriosis moreover varies between subspecialists. It thereover seems logical to match judgment and expertise to judge evidence and to involve thereafter all stakeholders in the translation into clinical recommendations. It should also be realised that interpretation and judgment vary with the understanding of pathophysiology. Whether endometriosis is seen as one or several G-E different diseases will help to understand that some 50% of typical lesions are not painful and that response to medical therapy is absent or inadequate in 10% to 40% respectively (56, 70). Future trials should reflect this and

it seems logical that at least superficial, cystic ovarian and deep endometriosis are evaluated and reported separately being clinically different entities (71).

Clinicians appreciate the achievements of EBM but were educated and did grow up with significances and P-values, and thus risk having misused their limited value to confirm a hypothesis (1). It was refreshing to realise the importance of the prior hypotheses (18) and to understand that clinical medicine has a Bayesian approach. Seeing a woman of that age, with these antecedents and these symptoms, results in many differential diagnoses which are refined into a workable probability by additional exams and tests to result in a treatment considering the consequences of mistakes and complications. Clinical medicine is also highly multivariate, with independent and dependent variables. Unfortunately, the gap between statistical inference and clinical understanding seems to be widening as illustrated in a recent report describing a new diagnostic test using a 'penalized regression model and machine learning with random forest' (72). This risks not being readily understood by most clinicians.

The quality and ranking of evidence need to be re-evaluated for medical therapy of endometriosis. The judgement is moreover bound to change if peritoneal fluid concentrations and progesterone resistance are taken into account (70, 73) and if endometriosis lesions are no longer considered a homogeneous group (19, 74) as illustrated by the biochemical heterogeneity (55) and by some deep lesions that continue to grow during medical therapy (56) or after menopause (75). Besides adjusting statistical inference, many aspects need to be defined such as "adequate" pain relief to continue treatment, or placebo effect without blinding, or 'women with proven endometriosis' after laparoscopy with surgical treatment. Nevertheless, we think that the clinical treatment of superficial endometriosis could be summarised as follows. Women with proven or suspected endometriosis and pain deserve a trial with medical therapy. However, the eventual growth of lesions during therapy should be monitored and if pain relief is inadequate, other options should be considered.

The quality and completeness of surgery for endometriosis are poorly defined. The severity of the disease and the surgery are variable, with cystic and deep endometriosis being technically difficult and complication prone causing oocyte damage, sexual problems and bladder, ureteral and bowel complications (76). Randomisation is irrealistic and can be unethical when surgeons are not equally trained in the techniques to be compared. However, It is suggested that the technique, results and complications can be judged by the small group of deep endometriosis surgeons with a large experience over a longer period. They know each other and understand the surgery each of them is doing, and the aspects they agree upon are probably high ranking evidence. This is not contradicted by the decision of doing a bowel resection or a conservative excision or a discoid excision being based to a large extent on personal preferences (77) since results and complications vary with surgical skills and experience. To convert this 'consensus opinion of experts' into evidence is a methodological and statistical challenge for the future.

In conclusion, an EBM approach to endometriosis has specific challenges. The diagnosis is limited to those undergoing laparoscopy and this clinical decision is based on a variable mixture of clinical exams and symptoms and imaging. The accuracies of imaging as ultrasound or MRI are well described (78, 79), but the predictive values vary with the locally variable prevalences, and their importance in clinical decision making varies from little (58, 76) to very much (80). Not only the indication for laparoscopy is variable also the recognition of endometriosis (81) as demonstrated for subtle lesions and observed for deep and appendiceal endometriosis. Medical therapy needs re-appraisal and for extensive surgery, the judgement of surgeons with experience needs validation. The complexity will need better integration of statistical analysis and inference to understand which exams and therapies improve outcomes (82).

Acknowledgements.

We do thank Manu Lesaffre and Geert Page, Belgium and Lone Hummelsjoy, United kingdom for fruitful discussions.

Funding

No funding

Disclosure of interests

None of the authors has anything to disclose

Contribution of authors

All authors have discussed endometriosis diagnosis and especially surgery during several yearly meetings and live surgeries. These discussions moreover resulted in a series of joint publications on endometriosis. PK's interest in statistics initiated this manuscript with the help of ML author of "Bayesian biostatistics". The text was subsequently edited and discussed by all authors.

Ethics

Ethical approval was not necessary

List of legends

Fig 1. Evidence in medicine starts with observations and trials, which have less risk of bias. Traditional statistics test the Null hypotheses resulting in P-values. Bayesian analysis is better suited to judge the hypothesis. Clinical experience is important for each aspect and judges the risk of bias provides the prior information to perform trials, orient analysis and evaluate external validity and grades of evidence. Considering the importance of clinical experience, the variability of experience by the sub-disciplines in endometriosis needs to be formally addressed to understand diagnosis and therapy.

Reference List

1. Nuzzo R. Statistical errors. Nature. 2014;506(7487):150.

2. Sellke T, Bayarri M, Berger JO. Calibration of ρ values for testing precise null hypotheses. The American Statistician. 2001;55(1):62-71.

3. Goodman SN. Of P-values and Bayes: a modest proposal. Epidemiology. 2001;12(3):295-7.

4. Goodman S, Greenland S. Why most published research findings are false: problems in the analysis. PLoS medicine. 2007;4(4):e168.

5. Ioannidis JP. Why most published research findings are false. PLoS medicine. 2005;2(8):e124.

6. Boos DD, Stefanski LA. P-Value Precision and Reproducibility. The American Statistician. 2011;65(4):213-21.

7. Colquhoun D. The reproducibility of research and the misinterpretation of p -values. Royal Society Open Science. 2017;4(12):171085.

8. Wasserstein RL, Lazar NA. The ASA Statement on p-Values: Context, Process, and Purpose, , 70:2, 129-133,. The American Statistician. 2016;70(2):129-33.

9. Hardwicke TE, Goodman SN. How often do leading biomedical journals use statistical experts to evaluate statistical methods? The results of a survey. PLOS ONE. 2020;15(10):e0239598.

10. Wilson JR, Falcone T. The JMIG Issues New Guidelines on Statistical Reporting and p-values. Journal of Minimally Invasive Gynecology. 2020;27(1):1-3.

11. Price R, Bethune R, Massey L. Problem with p values: why p values do not tell you if your treatment is likely to work. Postgraduate Medical Journal. 2020;96(1131):1-3.

12. Djulbegovic B, Guyatt GH. Progress in evidence-based medicine: a quarter century on. Lancet. 2017;390(10092):415-23.

13. Lawrence R, Force USPST. Guide to clinical preventive services: DIANE publishing; 1989.

14. Medicine CfE-B. Oxford Centre for Evidence-based Medicine – Levels of Evidence2009

15. Murad MH, Asi N, Alsawas M, Alahdab F. New evidence pyramid. Evidence Based Medicine. 2016;21(4):125-7.

16. Sackett DL, Rosenberg WMC, Gray JAM, Haynes RB, Richardson WS. Evidence based medicine: what it is and what it isn't. BMJ. 1996;312(7023):71-2.

17. Connor J. How to lie with statistics. LLC: Unlimited Press Works 2015.

18. Lesaffre E, Lawson AB. Bayesian biostatistics. United Kingdom: John Wiley & Sons, ltd; 2012.

19. Koninckx PR, Ussia A, Adamyan L, Wattiez A, Gomel V, Martin DC. Heterogeneity of endometriosis lesions requires individualisation of diagnosis and treatment and a different approach to research and evidence based medicine. Facts Views Vis Obgyn. 2020;11(3):263.

20. Devito NJ, Goldacre B. Catalogue of bias: publication bias. BMJ Evidence-Based Medicine. 2019;24(2):53-4.

21. Lin L, Chu H. Quantifying publication bias in meta-analysis. Biometrics. 2018;74(3):785-94.

22. Drossman DA, Whitehead WE, Camilleri M. Irritable bowel syndrome: a technical review for practice guideline development. Gastroenterology. 1997;112(6):2120-37.

23. Pepe MS. Combining diagnostic test results to increase accuracy. Biostatistics. 2000;1(2):123-40.

24. Nisenblat V, Prentice L, Bossuyt PM, Farquhar C, Hull ML, Johnson N. Combination of the non-invasive tests for the diagnosis of endometriosis. Cochrane Database Syst Rev. 2016;7:CD012281.

25. Chen Z, Hwang BS, Kim S. A correlated Bayesian rank likelihood approach to multiple ROC curves for endometriosis. Statistics in Medicine. 2019;38(8):1374-85.

26. Chen Z, Hwang BS. A Bayesian semiparametric approach to correlated ROC surfaces with stochastic order constraints. Biometrics. 2019;75(2):539-50.

27. Polin HB, Plaut ME. Chloramphenicol. N Y State J Med. 1977;77(3):378-81.

28. Murad MH, Altayar O, Bennett M, Wei JC, Claus PL, Asi N, et al. Using GRADE for evaluating the quality of evidence in hyperbaric oxygen therapy clarifies evidence limitations. J Clin Epidemiol. 2014;67(1):65-72.

29. Mahmood S, Nona P, Villablanca P, Nunez-Gil I, Ramakrishna H. The Meta-Analysis in Evidence-Based Medicine: High-Quality Research When Properly Performed. Journal of Cardiothoracic and Vascular Anesthesia. 2021;35(9):2556-8.

30. Ioannidis JP. The Mass Production of Redundant, Misleading, and Conflicted Systematic Reviews and Meta-analyses. Milbank Q. 2016;94(3):485-514.

31. Whiting PF, Rutjes AW, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. Ann Intern Med. 2011;155(8):529-36.

32. Yang B, Mallett S, Takwoingi Y, Davenport CF, Hyde CJ, Whiting PF, et al. QUADAS-C: A Tool for Assessing Risk of Bias in Comparative Diagnostic Accuracy Studies. Ann Intern Med. 2021;174(11):1592-9.

33. You SC, Krumholz HM. The Evolution of Evidence-Based Medicine: When the Magic of the Randomized Clinical Trial Meets Real-World Data. Circulation. 2022;145(2):107-9.

34. Letterie G. Three ways of knowing: the integration of clinical expertise, evidence-based medicine, and artificial intelligence in assisted reproductive technologies. Journal of Assisted Reproduction and Genetics. 2021;38(7):1617-25.

35. Murad MH, Asi N, Alsawas M, Alahdab F. New evidence pyramid. Evid Based Med. 2016;21(4):125-7.

36. Howick J. Exploring the Asymmetrical Relationship Between the Power of Finance Bias and Evidence. Perspect Biol Med. 2019;62(1):159-87.

37. Klemperer D. Drug Research. Deutsches Aerzteblatt Online. 2010.

38. Gaeta R, Gentile N. Evidence, discovery and justification: the case of evidence-based medicine. J Eval Clin Pract. 2016;22(4):550-7.

39. Allen R, Cai AG, Tepler P, deSouza IS. The "NUTS" statistic: Applying an EBM disease model to defensive medicine. J Healthc Risk Manag. 2021.

40. Kempers RD, Dockerty MB, Hunt AB, Symmonds RE. Significant postmenopausal endometriosis. Surg Gynecol Obstet. 1960;111:348-56.

41. Koninckx PR, Ussia A, Zupi E, Gomel V. Association of Endometriosis and Adenomyosis: Vast Literature but Scant Conclusive Data. J Minim Invasive Gynecol. 2018;25(5):745-8.

42. Koninckx PR, Ussia A, Adamyan L, Tahlak M, Keckstein J, Wattiez A, et al. The epidemiology of endometriosis is poorly known as the pathophysiology and diagnosis are unclear. Best Pract Res Clin Obstet Gynaecol. 2021;71:14-26.

43. Goodman LR, Franasiak JM. Efforts to redefine endometriosis prevalence in low-risk patients. BJOG. 2018;125(1):63.

44. Van Holsbeke C, Van Calster B, Guerriero S, Savelli L, Leone F, Fischerova D, et al. Imaging in gynaecology: How good are we in identifying endometriomas? Facts Views Vis Obgyn. 2009;1(1):7-17.

45. Demco L. Review of pain associated with minimal endometriosis. JSLS. 2000;4(1):5-9.

46. 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(9):2017-23.

47. Marcoux S, Maheux R, Berube S. Laparoscopic surgery in infertile women with minimal or mild endometriosis. Canadian Collaborative Group on Endometriosis [see comments]. N Engl J Med. 1997;337(4):217-22.

48. Jacobson TZ, Duffy JM, Barlow DH, Koninckx PR, Garry R. WITHDRAWN: Laparoscopic surgery for pelvic pain associated with endometriosis. Cochrane Database Syst Rev. 2014(8):CD001300.

49. Muzii L, Miller CE. The singer, not the song. J Minim Invasive Gynecol. 2011;18(5):666-7.

50. Ramirez PT, Frumovitz M, Pareja R, Lopez A, Vieira M, Ribeiro R, et al. Minimally Invasive versus Abdominal Radical Hysterectomy for Cervical Cancer. N Engl J Med. 2018;379(20):1895-904.

51. Perrotta M, Geampana A. Enacting evidence-based medicine in fertility care: Tensions between commercialisation and knowledge standardisation. Sociology of Health & Illness. 2021;43(9):2015-30.

52. Sampson JA. Peritoneal endometriosis due to the menstrual dissemination of endometrial tissue into the peritoneal cavity. Am J Obstet Gynecol. 1927;14:422-69.

53. Sampson JA. Perforating Hemorrhagic (chocolate)cysts of the ovary Trans Am Gynecol Soc 1921;45:162-241.

54. Koninckx PR, Martin DC, Donnez J. Do we need to separate initiation and growth to understand endometriosis? Fertil Steril. 2020;114(4):766-7.

55. Bulun SE, Yilmaz BD, Sison C, Miyazaki K, Bernardi L, Liu S, et al. Endometriosis. Endocrine reviews. 2019;40(4):1048-79.

56. Vercellini P, Sergenti G, Buggio L, Frattaruolo MP, Dridi D, Berlanda N. Advances in the medical management of bowel endometriosis. Best Practice & Research Clinical Obstetrics & Gynaecology. 2021;71:78-99.

57. Becker CM, Gattrell WT, Gude K, Singh SS. Reevaluating response and failure of medical treatment of endometriosis: a systematic review. Fertil Steril. 2017;108(1):125-36.

58. Koninckx PR, Ussia A, Wattiez A, Adamyan L, Martin DC, Gordts S. The severity and frequency distribution of endometriosis subtypes at different ages: a model to understand the natural history of endometriosis based on single centre/single surgeon data. Facts, Views and Vision in ObGyn. 2021;13(3):211-21.

59. Langweiler MJ. Evidence-Based Medicine and the Potential for Inclusion of Non-Biomedical Health Systems: The Case for Taijiquan. Front Sociol. 2020;5:618167.

60. Afrin S, AlAshqar A, El Sabeh M, Miyashita-Ishiwata M, Reschke L, Brennan JT, et al. Diet and Nutrition in Gynecological Disorders: A Focus on Clinical Studies. Nutrients. 2021;13(6).

61. Harris HR, Eke AC, Chavarro JE, Missmer SA. Fruit and vegetable consumption and risk of endometriosis. Hum Reprod. 2018;33:715-27.

62. Halpern G, Schor E, Kopelman A. Nutritional aspects related to endometriosis. Revista da Associação Médica Brasileira. 2015;61(6):519-23.

63. Mira TAA, Buen MM, Borges MG, Yela DA, Benetti-Pinto CL. Systematic review and meta-analysis of complementary treatments for women with symptomatic endometriosis. Int J Gynaecol Obstet. 2018.

64. Della Corte L, Noventa M, Ciebiera M, Magliarditi M, Sleiman Z, Karaman E, et al. Phytotherapy in endometriosis: an up-to-date review. J Complement Integr Med. 2020;17(3).

65. Porpora MG, Brunelli R, Costa G, Imperiale L, Krasnowska EK, Lundeberg T, et al. A promise in the treatment of endometriosis: an observational cohort study on ovarian endometrioma reduction by N-acetylcysteine. Evid Based Complement Alternat Med. 2013;2013:240702.

66. Nirgianakis K, Egger K, Kalaitzopoulos DR, Lanz S, Bally L, Mueller MD. Effectiveness of Dietary Interventions in the Treatment of Endometriosis: a Systematic Review. Reprod Sci. 2021:ahead of print.

67. Peng J, Wang R, Ding Z, Song X. Effect assessment of laparoscopy in combination with traditional Chinese medicine decoction in the treatment of endometriosis: A protocol for systematic review and meta-analysis. Medicine (Baltimore). 2021;100(29):e26699.

68. Koninckx PR, Ussia A, Tahlak M, Adamyan L, Wattiez A, Martin DC, et al. Infection as a potential cofactor in the genetic-epigenetic pathophysiology of endometriosis: a systematic review. Facts Views Vis Obgyn. 2019;11(3):209-16.

69. Donnez O, Donnez J. Deep endometriosis: The place of laparoscopic shaving. Best Pract Res Clin Obstet Gynaecol. 2021;71:100-13.

70. Donnez J, Dolmans MM. Endometriosis and Medical Therapy: From Progestogens to Progesterone Resistance to GnRH Antagonists: A Review. J Clin Med. 2021;10(5).

71. Donnez J, Nisolle M, Smoes P, Gillet N, Beguin S, Casanas-Roux F. Peritoneal endometriosis and "endometriotic" nodules of the rectovaginal septum are two different entities. Fertil Steril. 1996;66(3):362-8.

72. Moustafa S, Burn M, Mamillapalli R, Nematian S, Flores V, Taylor HS. Accurate diagnosis of endometriosis using serum microRNAs. Am J Obstet Gynecol. 2020.

73. Koninckx PR, Ussia A, Adamyan L, Gomel V, Martin DC. Peritoneal fluid progesterone and progesterone resistance in superficial endometriosis lesions. Human Reproduction. 2021.

74. Donnez J. The heterogeneity of endometriotic lesions could be explained by their progesterone resistance. Hum Reprod. 2021;36:2624-25.

75. de Almeida Asencio F, Ribeiro HA, Ribeiro PA, Malzoni M, Adamyan L, Ussia A, et al. Symptomatic endometriosis developing several years after menopause in the absence of increased circulating estrogen concentrations: a systematic review and seven case reports. Gynecological Surgery. 2019;16(1):3.

76. Koninckx PR, Ussia A, Keckstein J, Malzoni M, Adamyan L, Wattiez A. Review on endometriosis surgery. Gynecology and Pelvic Medicine. 2021;4:38-.

77. Koninckx PR, Ussia A, Adamian L, Alsuwaidi S, Amro B, Gharbi H, et al. Conservative Surgery of Deep Bowel Endometriosis. In: Ferrero S, Ceccaroni M, editors. Clinical Management of Bowel Endometriosis: Springer nature Swizerland; 2020. p. 119-34.

78. Guerriero S, Pascual M, Ajossa S, Neri M, Musa E, Graupera B, et al. Artificial intelligence (AI) in the detection of rectosigmoid deep endometriosis. Eur J Obstet Gynecol Reprod Biol. 2021;261:29-33.

79. Guerriero S, Martinez L, Gomez I, Pascual MA, Ajossa S, Pagliuca M, et al. Diagnostic accuracy of sonography for detecting parametrial involvement in women with pelvic endometriosis: a systematic review and meta-analysis. Ultrasound Obstet Gynecol. 2021.

80. Malzoni M, Casarella L, Coppola M, Falcone F, Iuzzolino D, Rasile M, et al. Preoperative Ultrasound Indications Determine Excision Technique for Bowel Surgery for Deep Infiltrating Endometriosis: A Single, High-Volume Center. Journal of Minimally Invasive Gynecology. 2020;27(5):1141-7.

81. Taylor HS, Adamson GD, Diamond MP, Goldstein SR, Horne AW, Missmer SA, et al. An evidence-based approach to assessing surgical versus clinical diagnosis of symptomatic endometriosis. Int J Gynaecol Obstet. 2018.

82. Bernstein E, Wang TY. Point-of-Care Ultrasonography. JAMA Internal Medicine. 2021;181(12):1558.