EVIDENCE-BASED MEDICINE CORNER

EBM in action: "Does ovulation induction increase the risk of ovarian cancer?"

Abdelhamid Attia, M.D.

Professor of Obstetrics & Gynecology, Cairo University; President; Arab Federation of evidence-Based Medicine

Induction of ovulation is an every-day practice for all gynecologists working in the area of reproductive endocrinology. Women suffering from anovulatory, tubal, or idiopathic infertility are commonly exposed to induction of ovulation medications that are sometimes administered at high therapeutic doses for extended periods of times. However, in 1992, a pooled analysis of 12 case-control studies has pointed out to a 2.7-fold increased risk of ovarian cancer in infertile women treated with fertility drugs and a 27-fold increased risk if the woman had used fertility drugs and had never been pregnant before[1]. However, the risk was not increased in infertile women who had not undergone induction of ovulation compared with the general population[1]. This has thrown clinicians as well as patients in a dilemma that needs to be resolved.

Evidence-Based Medicine is defined as the conscientious, judicious, and explicit use of current best evidence in making decisions about the care of individual patients[2]. Evidence for harm of therapeutic interventions is commonly sought from observational studies namely cohort and case control studies. But knowing the liability of case-control studies to many types of major bias and knowing the controversy about the methodological quality of the analyzed studies, one should ask "Is there better evidence that confirm or negate this evidence?"

———————

Correspondences: Dr. Abdelhamid Attia, 18 El-Ghaith St., El-Agouza, Cairo, Egypt. Email: aattia@thewayout.net

To answer this question we have to follow the systematic approach of EBM. This approach requires transforming the clinical problem into an answerable PICO question, searching the literature for studies on the topic, appraising these studies then applying their results, if valid, into one's practice.

I- Formulating an answerable PICO question

The acronym PICO stands for: "P" is the problem of the patient, "I" is the intervention (or exposure) we want to ask about, "C" is the comparison intervention and "O" is the outcome we are looking for. So the question would be phrased as follows: In infertile women (P); does induction of ovulation (I) compared to no induction (C) increase ovarian cancer risk (O)?

II- Searching for studies addressing the PICO question

The next step is to search for studies that addressed this question. Following the hierarchy of evidence-based medicine; the best evidence is obtained from systematic reviews and meta-analysis followed by randomized controlled trials (RCTs), cohort studies, case control studies, then case series and case reports.

Search needs key words and key words are extracted from the formulated PICO question. The most important key words in this scenario are:

ovulation induction and ovarian cancer. However, to do a complete search we have also to consider all the synonyms of the extracted key words.

It is to be noted that all the searches done below were conducted on 30/6/2006.

1. Searching for a meta-analysis in the Cochrane Library (CLIB):

The first site to search is the Cochrane Library (http://www.thecochranelibrary.com) as it has the highest quality systematic reviews and meta-analysis. If we find an answer to a PICO question in a Cochrane systematic review we usually do not search any where else. Key words used for the search were ovulation induction and cancer. However, the search revealed no Cochrane systematic reviews on the topic.

2. Searching in Pubmed:

Pubmed (http://www.pubmed.gov) is the National Library of Medicine premier bibliographic database that contains bibliographic citations and author abstracts from more than 4,800 biomedical journals published in the United States and 70 other countries. The database contains over 14 million citations dating back to the mid 1960s. Coverage is worldwide, but most records are from English-language sources or have English abstracts.

Each article in the Pubmed is indexed under different Medical Subject Headings known as MESH terms. Searching Pubmed using the MESH terms makes the search more specific.
a. Searching for meta-analysis: Key words used were Ovulation Induction[MeSH] AND Ovarian Neoplasms[MeSH]. Using the limits tab at Pubmed we limit the search to meta-analysis. But this search retrieved no articles.
b. Searching for Randomized Controlled Trials: RCTs give the second best evidence after meta-analysis. But RCTs are used mainly to test therapeutic interventions for their benefits and not used primarily to test for harm questions. However, we may find a RCT that report about a harmful effect of a certain therapeutic intervention if this harm was detected occasionally during studying its therapeutic benefit. In this scenario Key words used were Ovulation Induction[MeSH] AND Ovarian Neoplasms[MeSH] and limiting the search to RCTs retrieved no articles.
c. Searching for cohort studies: The Pubmed does not have the function of limiting the search to cohort studies. So search for cohort studies needs a wider search without limits and from the abstracts you can pick cohort studies to evaluate. Since cohort studies, if well conducted, represent the next best source of evidence after RCTs and their evidence is better than that obtained from case-control studies, we will pick one of the best cohort studies retrieved to evaluate.

The most recent and second largest cohort study with 12193 participants and the largest number of ovarian cancers detected is that of Brinton et al conducted in 2004[3]. Owing to the large number of participants and cancers detected we are going to critically appraise this study to know whether we can use its results in our own practice or not.

III- Critical appraisal of the cohort study

To critically appraise an article about harm we have to examine its relevance, validity and the applicability of its results[4].

A- Relevance

Before you go into details of methodology and results, you have to assess whether the study is relevant to your patients and your practice or not.
Is your patient so different from those in the study that its results don't apply?
For results of a study to apply to your patients the study population should be similar to the patients you see commonly in your practice. The authors stated that the population of the study was formed of patients with primary or secondary infertility with a median age of 30 years and having endometriosis, anovulation, tubal disease/pelvic adhesions, male factor, cervical disorders, or uterine disorders. Interventions studied were clomiphene citrate (CC) and a variety of human gonadotropins (HMG), namely, Pergonal (Serono, Rockland, MA), Humegon (Organon, West Orange, NJ), and Metrodin (Serono, Rockland, MA).

B- Validity

Validity is the methodological quality of the research. It answers the question "does the research have sound methodology so that we can trust its results?" To assess validity a series of questions should be fulfilled.

1- Were there clearly defined groups of patients, similar in all important ways other than exposure to the treatment or other cause?

For a cohort study to give relevant results; exposed and unexposed groups should be similar in all criteria except the exposure studied (here exposure is induction of ovulation).

The cohort consisted of 12193 infertile women treated in one of five endocrinology clinics in USA between 1965 and 1988. This cohort has been compared, regarding risk of cancer, in several ways; total cohort to general population, CC users versus non users, HMG users versus non users and users of either CC or HMG versus non users of either.

The first analysis in the study assessed ovarian cancer risk in the whole cohort including patients who had induction of ovulation and those who had not. The authors calculated the standardized incidence ratio (SIR) by dividing the number of observed cancers in the whole cohort by the expected numbers in the general US population based on age, race, and calendar year-specific incidence disease rates for females. This type of analysis, comparing the incidence in the whole cohort to that of the total population instead of a concurrent control group of normal fertile women has many flaws. First, it doesn't separate the effect of infertility as a risk factor from that of induction of ovulation. Second, it introduces bias to this analysis rendering it more of a case series rather than a true cohort study. However, we are concerned more with the second analysis that compares cancer risk in women who underwent induction of ovulation to those who did not.

The second analysis assessed ovarian cancer risk according to medication usage within the cohort. However, there was overlap between the groups as some patients used both CC and HMG so their results was calculated in both groups one time

as a CC user and another time as a HMG user. Meanwhile although the authors stated that they have adjusted for ovarian cancer predictors and potential confounders between the groups, the article lacked a table comparing the basic characteristics of the different groups. So serious selection bias can not be excluded in this study as there may be differences in characteristics of the groups regarding the age, duration and cause of infertility, gravidity, breast feeding, age of menarche and menopause as well as any other risk factor or confounder for ovarian malignancy.

2- Were treatments/exposures and clinical outcomes measured in the same ways in both groups?

The retrospective nature of the study did not allow for uniform measurement of exposures and outcomes. Information about treatment (exposure) was incomplete. Data were obtained from records of the five centers participating in the study. This does not allow for complete information about medications used outside the 5 centers till the end of the study that might affect the outcome. Also determination of the occurrence of cancer varied between individuals. Of the 8432 patient whose cancer status was assessed; data for 5597 were obtained from a questionnaire, 2347 from cancer registries, 216 from clinic records, and 272 from National Death Index. Also diagnosis of cancer was made by record linkage and self-reports and although the authors tried to confirm the diagnosis from pathological hospital records they did not state the number of cases verified by pathology and those who were not. This approach also does not give a chance to correlate infertility treatment to a specific pathological diagnosis. In cases that did not return the questionnaire and those whose last known address was incorrect, the authors might have missed the true identification of cancers among these subjects and incorrectly assigned person-years until the end of the study.

3- Was the follow-up of study patients complete and long enough?

Of 12193 eligible patients 2442 (20%) were lost to follow-up and 1319 (10.8%) refused to

participate in the study. This represents more than 30% of the original cohort which is considered a major source of bias as the outcome in these patients may dramatically change the results of the study.

Another important point is that although the median length of follow-up among subjects was 18.8 years, with more than 80% followed for 15 or more years, the median age in the study was 30 years which makes the median age at end of follow-up 48.8 years. Terminating the follow-up at this age is not enough to study the incidence of an outcome whose peak age is the early sixties.

4- Cause and effect relationship

For a study to be valid, an important aspect is that there should be a cause and effect relationship between the exposure and the outcome that can be justified. This is examined in the following section.
a. Is it clear that the exposure preceded the onset of the outcome? In all cases induction of ovulation preceded the development of cancer and the investigators excluded the 3 cases that developed cancer ovary within the first year of follow-up as ovarian cancer has a relatively long latency period before its diagnosis. However, the time period between induction of ovulation and diagnosis of cancer was not clearly stated.
b. Is there a dose-response gradient? Although higher doses of CC appeared unrelated to risk (e.g., rate ratio = 0.80 for 2251 mg or more CC). However, there was some evidence of a slightly elevated risk, though statistically non significant, for women with either 12 or more cycles of exposure (rate ratio = 1.54, 95% CI 0.5, 5.1) or 15 or more years of follow-up (rate ratio = 1.48; 0.7, 3.2). Both of these risks, however, were based on few diagnosed cancers (3 and 5, respectively).
c. Is there positive evidence from a "dechallenge-rechallenge" study? Dechallenge-rechallenge evidence does not apply to this study as once the patient is exposed, the effect of exposure cannot be removed nor can the outcome revert to baseline to observe whether the event is reinitiated with subsequent exposure.
d. Is the association consistent from study to study? Results of this study are consistent with other cohort studies.

e. Does the association make biological sense? One of the mechanisms that support a possible cause-and-effect relationship, between the effect of ovulation induction on the ovary and later development of ovarian cancer, is the incessant ovulation theory postulated by Fathalla in 19715. He proposed that repeated ovulation exposes the ovarian epithelium to microtrauma that may lead to mitotic abnormalities and subsequently cancer. The decreased risk of cancer in the presence of prolonged anovulatory intervals, as in repeated pregnancies and oral contraceptive use, may also support this theory.

C- Results

1- What are the results?

The overall results of this study shows that infertility is associated with a statistically significant higher risk of ovarian cancer compared to general population (SIR 1.98; 95% CI 1.4, 2.6). However, data about induction of ovulation in infertile women is reassuring as it was not associated with an increased risk of ovarian cancer compared to those who do not undergo induction of ovulation. This finding was consistent for the type of drug used, dose, duration of use as well as the duration of follow-up.

2- Are the results of this study important?

In light of the inability to conduct RCTs to assess harm questions, cohort studies are the best design to investigate such a hypothesis. But although the investigators did their best using statistical adjustment and alternative analysis to assess the credibility of the results, still there are many limitations to the current study that prevent accepting its results as the best evidence for practice. These limitations are:
a. High rate of drop outs and loss to follow-up (30.8%): These very high rates of drop outs and loss to follow-up may have lead to a serious selection bias that could have affected the study results.
b. Incomplete data: the authors stated that 41% of the patients who were located as alive did not complete the questionnaire and they had to search clinic records and cancer registries to obtain their

data. Also a number of women had incomplete workups and data about drugs used for stimulation were less than optimal as the authors could not account for drugs subsequently prescribed by other providers among the women who did not complete the questionnaire.

c. Very low event rates: The total patients who were found to have cancer were 45 cases. Although this is the highest number found in a cohort study, still the high rate of drop out & loss to follow-up in comparison to the event rate may seriously affect the results of this study.

d. Relatively short follow-up period: As stated before, follow-up was not enough to reach the age of the peak incidence of ovarian cancer development. The need for a longer follow-up is also evident in the relatively elevated risk, though not-significant, in patients who received CC and were followed up for more than 15 years (rate ratio = 1.48) compared to those who were followed up for less than 15 years (rate ratio = 0.47) and for patients who received HMG and were followed up for more than 15 years (rate ratio = 2.46) compared to those who were followed up for less than 15 years (rate ratio = 0.67). This relatively short follow-up period is associated with a lesser incidence of ovarian cancer among infertile women in general therefore it would be difficult to assess a true difference between exposed versus unexposed women.

IV: Clinical resolution and implications on practice

Generally, results of this study are reassuring. However, given its limitations and problems in its methodology highlighted above, there is a need for a better study with more complete data and longer follow-up to an older age especially with the tendency of increased risk, albeit non significant, with higher doses and longer follow-up periods.

It is also to be noted that to date non of the studies published gives a final answer to the question: Does ovulation induction increase the risk of ovarian cancer?

REFERENCES

1. Whittemore AS, Harris R, Itnyre J, and the Collaborative Cancer Group. Characteristics relating to ovarian cancer risk -- collaborative analysis of 12 US case-control studies II: invasive epithelial cancers in white women. Am J Epidemiol 1992;136:1184-1203
2. Sackett DL, Rosenberg WMC, Gray JAM, Haynes RB, Richardson WS. Evidence based medicine: what it is and what it isn't. BMJ 1996;312:71-2.
3. Brinton LA, Lamb EJ, Moghissi KS, Scoccia B, Althuis MD, Mabie JE, Westhoff CL. Ovarian cancer risk after the use of ovulation-stimulating drugs. Obstet Gynecol. 2004;103:1194-1203
4. Levine M, Walter S, Lee H, Haines T, Holbrook A, Moyer V. Users' guides to the medical literature. IV. How to use an article about harm. Evidence-Based Medicine Working Group. JAMA. 1994;271:1615-1619
5. Fathalla MF. Incessant ovulation: a factor in ovarian neoplasia? [letter]. Lancet 1971;2:163