RE: Manuscript Number ONG-23-928

PREDICTING THE RESOLUTION AND PERSISTENCE OF OVARIAN CYSTS: FINDINGS FOR THE PRACTICING PHYSICIAN

Dear Dr. Pavlik:

Thank you for sending us your work for consideration for publication in Obstetrics & Gynecology. Your manuscript has been reviewed by the Editorial Board and by special expert referees. The Editors would like to invite you to submit a revised version for further consideration.

If you wish to revise your manuscript, please read the following comments submitted by the reviewers and Editors. Each point raised requires a response, by either revising your manuscript or making a clear argument as to why no revision is needed in the cover letter.

To facilitate our review, we prefer that the cover letter you submit with your revised manuscript include each reviewer and Editor comment below, followed by your response. That is, a point-by-point response is required to each of the EDITOR COMMENTS (if applicable), REVIEWER COMMENTS, and STATISTICAL EDITOR COMMENTS (if applicable) below. The revised manuscript should indicate the position of all changes made. Please use the "track changes" feature in your document (do not use strikethrough or underline formatting). Upload the tracked-changes version when you submit your revised manuscript.

Your submission will be maintained in active status for 21 days from the date of this letter. If we have not heard from you by 07/14/2023, we will assume you wish to withdraw the manuscript from further consideration.

EDITOR COMMENTS:

Please note the following:

1. Please shorten title.
2. This paper is of interest and relevance to the readers of this journal. It would benefit from being more streamlined and more concise (especially the Results section) so readers can more easily follow the results and understand clinical implications.
3. In current form, the results section is almost like a listing of numerous analyses which makes it challenging to follow.

* Help us reduce the number of queries we add to your manuscript after it is revised by reading the Revision Checklist at https://journals.lww.com/greenjournal/Documents/RevisionChecklist_Authors.pdf and making the applicable edits to your manuscript.

***Plagiarism - Include if CrossCheck report indicates a text match with a previously published source.***

* All submissions that are considered for potential publication are run through CrossCheck for originality. The following lines of text match too closely to previously published works or need to be cited:
  - Please rephrase lines 76-88, "Three scanning approaches...work reported here." Please rephrase Lines 88-96, "Thus, a visualization...estimates of resolution."
  - Was this also presented at the Markey Cancer Center Research Day on May 10, 2022? If so, please add this to your meetings information on the title page.

* Figures 1-3: Please remove the figures from your manuscript file and upload separately as individual figure files.

REVIEWER COMMENTS:

Reviewer #1:
Thank you for the opportunity to review the manuscript titled "Predicting the resolution and persistence of ovarian cysts: findings for the practicing physician."
The importance of this study is to report on the resolution of septated and unilocular cysts using the University of Kentucky Ovarian Cancer Screening Trial experience.

This is not a totally original idea in that other examples in the literature try to provide some data to support guidelines on a common clinical issue: How to follow a non suspicious ovarian cyst. This database is much larger and thorough then any other retrospective experience. One example in the literature is a conference consensus on how to follow simple adnexal cysts. Levine D. Simple adnexal cysts: SRU consensus on follow-up and reporting. Radiology. 2019. Another example, is this study which suggested it would take over a year in between scan to determine if a cyst in enlarging. Suh-Burgmann, Ovarian Cystadenomas: growth rate and reliability of imaging measurements. J Ultrasound Med 2022. PMID 34846072. The references on this submission do not include these studies but seem to list other examples from 2019 or earlier. This study is unique in the duration and size of prospectively collected data available for the authors analysis.

The abstract and precise is concise and well written. I wonder if simplifying this to understanding how long before a simple cyst resolves is sufficient since the differences in subgroups seem clinically limited. Is median time of 1 year versus 1.25 years clinically meaningful?

I wonder why the authors decided to use different # of categories for the multifactorial analysis. For example, what is the significance for the authors differentiating a 3-6 cm category. Why not less than or greater than 6 cm. Similarly, did the various BMI groups help with clarifying the analysis or would a simple greater than or less than 35 be sufficient?

Although I noted that the study collected data from 31 years, I could not find the follow up for the subgroup of patients who had incidence cyst resolved. For example, how many follow up ultrasounds (encounters) were done before and after resolution. I did notice a Median time to resolution. Also how many cases were eventually censored for surgery and after how long and how many ultrasounds?

I defer to statistician but wonder why include BMI and family history in multivariant analysis if not significant in univariant analysis?

The readability of the study would benefit from greater inclusion of case numbers instead of only percentages and number of ultrasounds(or encounters) to understand how patients were followed and how to apply this knowledge in clinical practice. For example, why do the authors believe 4x the median resolution time is recommended? If HRT used then need 8 years of follow up for a simple cyst? But we do not know how many patients during that follow up period of repeat scans decided to have surgery since they were censored.

This study tries to provide important guidelines for practitioners who commonly see non-suspicious cysts.

Other specific questions for the authors:

- Figure 1 staggering number of potential encounters 321,865. Can the authors clarify the time that this breakdown occurred? And can and additional table provide the patient characteristics for the cases included in this particular analysis. So how many family history patients were there (7996)?
- Figure 3 should be consistent with table 2. So KaplanMeier curves with <3, 3-6, >6 cm
- Table 1 In premenopausal women how do the authors propose to know that the incident cyst time to resolution are the same cyst and not another incident cyst? And can the legend explain something about the time frame or number of encounters per resolution.
- Table 2 If age by decade was not statistically significant why is age <70+? Also perhaps months instead of 0.984 years would be easier to read
- Line 21 how do the authors explain the finding that cysts resolved were younger but in the next sentence the fastest was 70+ (line 22)
- Line 35 perhaps a definition and/or etiology of ovarian cyst in introduction
- Line 59 what was the average or median follow up of participant in this retrospective analysis.
- Line 73 nice description of ultrasound assessment
- Line 95 how many censored had surgery? And after how many encounters?
- Line 98 when did many cysts develop after how many ultrasounds?
- Line 111 321,865 or 321, 566 encounters. And line 118 consider adding case number for ease of reading
- Line 130 define asynchronous vs synchronous
- Line 180 would this findings hold if analysis was just less than or greater than 6 cm?
- Line 196 please explain why four times the expected median is the recommendation of the authors. After how many years did ultrasounds stop? What is the percentage of patients that say resolution after 4 years or 6 years?
- Line 216 do the authors believe that the statistical significant findings in some categories such as HRT use or cyst type merit an alternate time course for follow up imaging.

Overall, this a thought provoking paper that highlights some of the limitations of ovarian cyst follow up. Particularly in premenopausal women this is challenging since cyst formation is multifactorial. I believe the readers of this journal appreciate this perspective but wonder about the clinical applicability of some of the presented data and wonder if the authors believe any follow up ultrasounds for simple cysts required.
Reviewer #2:
1. Precis - include all seven factors analyzed
2. Lines 25-26 states the following: "Individuals who reported a family history of ovarian cancer (64.3%) were more likely to resolve than those that did not (35.7%) (p<0.05). Menopausal status, BMI, or family history made no difference on resolution time (p>0.05). Is family history predictive?
3. How did you confirm family history of ovarian cancer and hormonal replacement therapy?
4. Line 68 - How did you determine menopausal status for patients s/p hysterectomy?
5. Line 71 - How many ultrasound technicians performed scans and were they all trained to perform scans using methodology described?
7. Line 196 - How did you conclude surveillance should continue four times beyond the expected median resolution time?

Reviewer #3:
Review of Manuscript ONG-23-928 "Predicting the resolution and persistence of ovarian cysts: Findings for the practicing physician"

A manuscript that utilizes data on TVUS results gathered from the University of Kentucky Ovarian Cancer screening study has been submitted. As noted by the authors, the objective of the study was to evaluate resolution of incident cysts that occurred during prospective surveillance. The authors subsequently performed analyses to determine the potential factors that may predict resolution as well as timing of said resolution. They have not included a STROBE checklist which would be expected with the reporting of cohort trial data. A minor point is that in the manuscript p values are listed both as p < .05 and p < 0.05, please revise to be consistent. I have the following questions and comments.

Précis - This seems to be an incomplete thought. Please revise.

Abstract - If space allows add the years to the abstract. Also note in the abstract these are results from a prospective cohort study.
Line 12 - minor issue but "t" for t-test should be in italics I believe.

Introduction - Line 50 - minor point but concern for malignancy may be present even if less than 10cm so perhaps remove the size limitation from this comment and/or split this into 2 reasons for surgery - concern for cancer and size > 10cm.

Methods - Line 62-6 - What about prior Gyn surgery? Were women with a hysterectomy eligible to enroll? How about a prior USO with no prior surgeries?
Line 65 - Is this individual BRCA status or family status in terms of a pathogenic variant?
Line 105 - See previous comments on t-test.

Results - Line 115 - please state that those with prevalent cysts were excluded from this analysis.
Line 134 - Can you provide a range or IQR for the resolution?
How often was surgery contemplated for these patients? Is there information on women that could serve as a different "control" that had persistence of benign cysts but that underwent surgery instead of continued TVUS assessments?

Discussion - Line 195 - So to be clear for your data with resolution of 14+ months the recommendation for continued surveillance would be 56+ months?
Line 201 - Did HRT predate the presence of a cyst?
Line 204 - So if HRT may promote resolution, why did HRT not prevent occurrence?

The tables and figure are duplicated

Tables - Consider making a supplemental table that provides summary information about these selected patients, what made them eligible, and baseline demographics.

Table 1 - It seems like much of this information is included in the manuscript. If so, perhaps supplemental.
Tables 2 and 3 - No comments

Figures - Figure 1 - Any information on how many of the 971 with persistent cysts had surgery?
Figures 2 and 3 - No comments
STATISTICAL EDITOR COMMENTS:

lines 22-23 and Table 2: In Table 2, the stats test by age strata was a test for an overall difference, not a specific comparison one age group. Need to provide CIs for each of the age groups (see below comments re: Table 2) and demonstrate whether the p < 0.05 difference was for age 70+ vs each of the other groups, or simply vs the < 40 age group.

Fig 2: This appears to be a K-M graph, but need to explicitly label it. Also, need to show the N at risk for each group at the indicated time points along the x-axis. As long as the censoring rates remain similar for the two groups vs time, then the estimates for % remaining or its converse % resolution is valid. If there is a differential rate of censoring, then the estimates are likely biased, since one is assuming an equivalence of outcomes for those shown on the K-M graphs vs those censored from the graphs.

Fig 3: Again, these appear to be K-M graphs, but they need to be labelled explicitly. Also, need to include either below the xx-axes at the indicated time points, or in a separate Table, the N for those remaining at risk for each group.

Table 1: Need to include CIs for each of the % resolved.

Table 2: Need to include IQRs for each of the median resolution times cited. Should also indicate the N for each subset.

Table 3: Should round all HRs and their CIs to 0.01 precision. Need to identify the referent group. Need to state whether the assumption of proportion hazard rates was examined and verified. Need units for BMI. Need to include N for each subset. Need to list the variables retained in the final multivariate models in footnote to Table.

--
Sincerely,
Vivian W. Sung, MD, MPH
Deputy Editor, Gynecology

The Editors of Obstetrics & Gynecology

In compliance with data protection regulations, you may request that we remove your personal registration details at any time. (Use the following URL: https://www.editorialmanager.com/ong/login.asp?a=r). Please contact the publication office if you have any questions.
Dr. Jason D. Wright  
Editor-In-Chief,  
*Obstetrics and Gynecology*

Dr. Wright,

I am submitting the revision of our paper titled, “*PREDICTING THE RESOLUTION AND PERSISTENCE OF OVARIAN CYSTS: FINDINGS FOR THE PRACTICING PHYSICIAN*” for publication in *Obstetrics and Gynecology*. In this work we have focused on the resolution of incident ovarian cysts in relation to diameter, structure, age, BMI, hormone replacement therapy (HRT), family history of ovarian cancer and menopausal status to help the practicing physician decide between surveillance or surgical intervention. This paper presents evidence that cysts <3 cm resolve faster than cysts >6 cm and that septated cysts resolve faster than unilocular cysts. Cyst resolution occurs faster in individuals over age 70. Individuals using HRT had longer cyst resolution times. In the revision we are able to provide a method that can be used to estimate the surveillance interval of ovarian cysts, providing clinicians with expectations for time to resolution in relation to cyst type, size, HRT use and age, allowing the practicing physician to implement a time-course for follow-up imaging.

We have responded to every point raised in review, making changes that were asked for. I have placed our responses to points raised in review at the bottom of this letter. I have used green highlighting to mark our responses to the original editors comments which are in font that is not highlighted.

I feel that this work is particularly suitable for publication in *Obstetrics and Gynecology* because it focuses on information that will be of value to the practicing physician. This work is significant, and within the scope of *Obstetrics and Gynecology* due to its focus on practical information. This work is an original submission to *Obstetrics and Gynecology* and is not under consideration for publication in another journal. All authors have approved the manuscript and agree with its submission to *Obstetrics and Gynecology*. I look forward to the appearance of our work in *Obstetrics and Gynecology*.

Sincerely, Edward Pavlik, Professor  
Division of Gynecologic Oncology,  
University of Kentucky College of Medicine

---

**EDITOR COMMENTS:**

Please note the following:
1. Please shorten title.
2. This paper is of interest and relevance to the readers of this journal. It would benefit from being more streamlined and more concise (especially the Results section) so readers can more easily follow the results and understand clinical implications.
3. In current form, the results section is almost like a listing of numerous analyses which makes it challenging to follow.

* Help us reduce the number of queries we add to your manuscript after it is revised by reading the Revision Checklist at [https://journals.lww.com/greenjournal/Documents/RevisionChecklist_Authors.pdf](https://journals.lww.com/greenjournal/Documents/RevisionChecklist_Authors.pdf) and making the applicable edits to your manuscript.

***Plagiarism - Include if CrossCheck report indicates a text match with a previously published source.***
* All submissions that are considered for potential publication are run through CrossCheck for originality. The following lines of text match too closely to previously published works or need to be cited:
  - Please rephrase lines 76-88 (now 86-99), "Three scanning approaches...work reported here."

**Lines 86-99 have been re-phrased and identified with green highlighting in the manuscript.**

Please rephrase Lines 88-96 (now 99-116), "Thus, a visualization...estimates of resolution."

**Lines 99-117 have been re-phrased and identified with green highlighting in the manuscript.**

- Was this also presented at the Markey Cancer Center Research Day on May 10, 2022? If so, please add this to your meetings information on the title page.

**This presentation was added to the title page**
* Figures 1-3: Please remove the figures from your manuscript file and upload separately as individual figure files.

The figures were removed from the manuscript and uploaded as individual figures.
Reviewer #1:
Thank you for the opportunity to review the manuscript titled "Predicting the resolution and persistence of ovarian cysts: findings for the practicing physician."

The importance of this study is to report on the resolution of septated and unilocular cysts using the University of Kentucky Ovarian Cancer Screening Trial experience.

This is not a totally original idea in that other examples in the literature try to provide some data to support guidelines on a common clinical issue: How to follow a non suspicious ovarian cyst. This database is much larger and thorough then any other retrospective experience. One example in the literature is a conference consensus on how to follow simple adnexal cysts. Levine D. Simple adnexal cysts: SRU consensus on follow-up and reporting. Radiology. 2019.

I am well aware of this paper (DOI: 10.1148/radiol.2019191354 since I am one of the authors (EJP). It is not included in the citations because it is an expert consensus and is not data driven.

Another example, is this study which suggested it would take over a year in between scan to determine if a cyst in enlarging. Suh-Burgmann, Ovarian Cystadenomas: growth rate and reliability of imaging measurements. J Ultrasound Med 2022. PMID 34846072.

The authors are aware of this publication (J Ultrasound Med. 2022 Sep;41(9):2157-2167, DOI: 10.1002/jum.15895), which deals chiefly with variation in measurements between two ultrasound observers. As such, it deals with measurement per se and not cyst life directly. For this reason, we choose not to include it in the references.

The references on this submission do not include these studies but seem to list other examples from 2019 or earlier. The DOI: 10.1002/jum.15895 study is not a prospective study, while the study under review is unique in the duration and size of prospectively collected data available for the authors analysis.

The abstract and precise is concise and well written. I wonder if simplifying this to understanding how long before a simple cyst resolves is sufficient since the differences in subgroups seem clinically limited. Is median time of 1 year versus 1.25 years clinically meaningful? The value of the work that is presented is that it should
make the clinician aware of the spectrum of resolution times, and that they vary with
certain co-factors. This awareness will not be served by a simplification that ignores
the co-factors. We emphasize in the revision (lines 7, 32-33, 226-238, 253-257) that
we are introducing a novel way of estimating the surveillance intervals relative to
various co-factors.

We feel that median measurements per se, as pointed out by this reviewer, may not
appear clinically meaningful, as expressed in the univariate analysis in Table 3.
However, only the mid-point of each effect is being reported in Table 2. In Table 4,
the entire effect is expressed in terms of the HR with comparison between univariate
and multivariate HR results. We believe that the meaningful clinical picture emerges
from Table 4.

I wonder why the authors decided to use different # of categories for the
multifactorial analysis. For example, what is the significance for the authors
differentiating a 3-6 cm category. Why not less than or greater than 6 cm. We made
this choice so that the individual contributions of small cysts could be observed by the
readers. The reviewer will remember that small cysts were a highly controversial
topic in the deliberations of the expert group reported in DOI:
10.1148/radiol.2019191354. Consequently, we feel that this group should not be
ignored. However, as now clarified in the legend to the modified Table 4, “Diameter,
Age, and BMI were analyzed as continuous variables for both univariate and
multivariate analysis. Cyst structure, HRT Use, Menopausal Status, and Family history
of Ovarian Cancer were analyzed as nominal variables.” Thus, this modification
avoids the reviewer’s criticism. Similarly, did the various BMI groups help with
clarifying the analysis or would a simple greater than or less than 35 be sufficient?
The issue that is central with regard to BMI is the possibility that ultrasonography can
underperform as women become larger. Consequently, we felt that utilizing the BMI
categories that could portray increasing size, would show the reader if there was an
impact of ultrasonographic underperformance as size increased. Since neither
univariate nor multivariate indicated any differences between the BMI ranges that
we utilized, we do not feel that underperformance is a factor in our determinations.
However, as now clarified in the legend to the modified Table 4, “Diameter, Age, and
BMI were analyzed as continuous variables for both univariate and multivariate
analysis.” Thus, this modification avoids the reviewer’s criticism.
Although I noted that the study collected data from 31 years, I could not find the follow up for the subgroup of patients who had incidence cyst resolved. We considered a cyst resolution event to be the time between the first ultrasonographic identification of a cyst and the first ultrasonographic absence of the cyst. Upon identification of a cyst, cysts were followed at ~3month intervals. Resolution was verified by an additional serial ultrasound at 3-6 months later and was the standard follow up that was utilized. For example, how many follow up ultrasounds (encounters) were done before and after resolution. The number of ultrasounds prior to cyst discovery was quite variable due to physiology/biology. We reported the exact duration of the cyst because that is the most meaningful to the reader, and every effort was made to perform re-evaluations at 3 month intervals from which for any time point the number of ultrasound encounters is estimable. We introduced the follow-up estimate in line 67. I did notice a Median time to resolution. Likewise, from the medians reported the 3 month interval can be used to estimate the number of ultrasound exams. Also how many cases were eventually censored for surgery and after how long and how many ultrasounds? Of the 47,757 women who received 634,566 ultrasounds, 2090 women representing 14,563 ultrasounds were censored from analyses because they received surgery. “For how long”: completely, starting from their first encounter going to their last encounter.

I defer to statistician but wonder why include BMI and family history in multivariant analysis if not significant in univariant analysis?

If some X1 does not significantly predict Y in a bivariate correlation, it is possible that when more predictors (X2, X3 etc) are considered in a multivariate analysis, X1 can evaluate as a statistically significant predictor. This can happen when a variable such as X2 “suppresses” the true relationship between X1 and Y and the true relationship between X1 and Y only emerges when X2 is statistically controlled in the multivariate analysis. For this reason, we included all variables in both univariate and multivariate analyses. In the present study, we were aware of the possibility that family history of ovarian cancer might be doubled linked to a family history of elevated BMI which would be possibly revealed in the multivariate analysis.

The readability of the study would benefit from greater inclusion of case numbers instead of only percentages and number of ultrasounds(or encounters) to understand how patients were followed and how to apply this knowledge in clinical
practice. **We have added numbers of participants to Figure 1.** For example, why do the authors believe 4x the median resolution time is recommended? **We have made the particular assumption that since the data reported here evaluate as normally distributed with regard to time, size, and BMI then the medians and the means will be equal. 99.7% of a normal distribution will be within 3 standard deviations. In a standardized normal distribution, the mean is zero and the standard deviation is 1, so that a multiplier greater than 3 should include virtually all of the distribution. By rounding to 4, we extrapolate that all of the distribution of resolving cysts can be covered.** **We have added this explanation to lines 226-238.** If HRT used then need 8 years of follow up for a simple cyst? **By our premise, the answer is “yes” (6-8 yrs, lines 226-238), but some will resolve sooner, and our estimates allow for all cysts to resolve.** But we do not know how many patients during that follow up period of repeat scans decided to have surgery since they were censored. **The data that is described here is dependent on the censoring so that estimation of natural cyst resolution would not be influenced by a surgically removed cyst being treated as a resolved cyst or by some contributions of the surgical healing process. For answering the question posed by this reviewer, a data set would need to be used involving a sufficient number of individuals who had surgery.**

This study tries to provide important guidelines for practitioners who commonly see non-suspicious cysts.

Other specific questions for the authors:
- Figure 1 staggering number of potential encounters 321,865. Can the authors clarify the time that this breakdown occurred?

**Lines 64-66 states, “47,762 individuals enrolled in the University of Kentucky Ovarian Cancer Screening Trial (UK-OCST) from November 1987 to March 2019 and were evaluated in 321,566 screening encounters.” The dates also have been stated in the legend for Figure 1.**

And can and additional table provide the patient characteristics for the cases included in this particular analysis. So how many family history patients were there (?996)?
We have introduced Table 1 with patient characteristics and re-numbered the original tables.

- Figure 3 should be consistent with table 2. So KaplanMeier curves with <3, 3-6, >6 cm

Because of the way that the univariate and multivariate analyses were conducted (nominal vs continuous), we have re-presented Table 3 now Table 4) so that there is now no inconsistency between now Tables 3 and 4.

- Table 1 In premenopausal women how do the authors propose to know that the incident cyst time to resolution are the same cyst and not another incident cyst?

Included in the original legend to the table was the text, “Individuals with multiple cyst findings were included so that all cyst findings exceeded the total number of individuals in each group A, B, C.” Thus, we felt that we did have some ability to discriminate “another incident cyst”. However, the reviewer’s point is well-taken that if cyst 1 rapidly resolves and a cyst 2 appears within a short period of time, it is possible that two cyst events may get interpreted as one under the intervals employed here. To accommodate this possibility, text has been added to the legend of what is now Table 2: “Findings are subject to the surveillance intervals described here.”

And can the legend explain something about the time frame or number of encounters per resolution. This table is the “ever table” summarizing cysts that either resolve or continue to persist at some time over the entire 30+ year period of evaluation. The information about time is reflected elsewhere in Table 3. In addition, as already explained, upon identification of a cyst, cysts were followed at ~3month intervals. Resolution was verified by an additional serial ultrasound at 3-6 months later and was the standard follow up that was utilized. The number of ultrasounds prior to cyst discovery was quite variable due to physiology/biology. We reported the exact duration of the cyst because that is the most meaningful to the reader, and every effort was made to perform re-evaluations at 3 month intervals from which for any time point the number of ultrasound encounters is estimable.

- Table 2 If age by decade was not statistically significant why is age <70+?
The statement in the paper referred to by the reviewer is: “When segregated by decade, resolution times of both unilocular and septated cysts in individuals 40-69 years old were not significantly different.” This statement in the text refers to the decades 40-49 vs 50-59 vs 60-69 as not being significantly different. As a group expressed in now Table 3, the 40-69 age group is significantly different from <40 and 70+ (70 and older). Text has been added to emphasize this distinction (line 182-184).

Also perhaps months instead of 0.984 years would be easier to read

The authors feel that with the longest resolution time being 3.5 years, the resolution time was best represented in years rather than months for easier comprehension.

- Line 21 how do the authors explain the finding that cysts resolved were younger but in the next sentence the fastest was 70+ (line 22)

The text that the reviewer addresses is “Individuals with resolved cysts were younger (53.2+0.3 yrs) than individuals with persisting cysts (57.3+0.5 yrs) (p <0.05). Resolution was fastest in individuals age 70+, and most prolonged in individuals age <40 (p<0.05).” The first sentence refers to the age of individuals with resolving cysts vs those with persisting cysts, while the second sentence refers to speed of resolution. These are two different parameters. In a physiological sense, the regularly changing hormonal environment in younger women may have an influence that is cystogenic, and cystostatic. A decreased influence due to hormonal cystostasis in older women could explain faster resolution. In addition, decreased ovarian volume with age may be cystocidal.

- Line 35 perhaps a definition and/or etiology of ovarian cyst in introduction

Line 36-41 now includes a definition/etiology. It reads “A benign ovarian cyst or benign cystic ovarian mass is a solid or fluid-filled sac or pocket (cyst) within or on the surface of an ovary. Ovarian cysts can originate from hormonal changes that affect the ovary, especially ovarian follicles, or from endometriosis or from infections. It is important to distinguish a benign ovarian cyst from an ovarian malignancy.”

- Line 59 what was the average or median follow up of participant in this retrospective analysis.
This is a prospective analysis and we have added the requested parameters at line 66-67: “for an average of 6.7 screening encounters/person with a mean follow up of 7.9 ± 0.04 years.”

- Line 73 nice description of ultrasound assessment
  
  **Thank you!**

- Line 95 how many censored had surgery? And after how many encounters?
  
  **The following has been added at lines 117-120:**
  
  There were a total of 1334 cases that were censored due to surgery. Surgery due to screening accounted for 543 cases and incidental surgeries unrelated to screening accounted for 791 cases. The average number of encounters prior to surgery of individuals censored due to surgery was 7.2 ± 0.2.

- Line 98 when did many cysts develop after how many ultrasounds? **Due to individuals having more than a single cyst event, we are unable to provide this estimate**

- Line 111 ?321,865 or 321,566 encounters. And line 118 consider adding case number for ease of reading
  
  **Line 139 has been updated to “321,566” to reflect the accurate number of encounters. We prefer expressing these findings in terms of cyst number in order make the reader aware that some individuals had multiple cysts.**

- Line 130 define asynchronous vs synchronous
  
  **We added a definition in Lines 125-127, that now reads, “Those who had bilateral cysts findings were further classified as to bilateral cysts that appeared at the same time (synchronous) and those bilateral cysts that appeared at different times (asynchronous).”**
- Line 180 would this findings hold if analysis was just less than or greater than 6 cm?

The expression on a cm basis in what is now line 176 and Table 3 would be less robust if the dichotomization inquired about by the reviewer were utilized. For example, this can be understood if global temperature were to be expressed as less than or greater than freezing.

- Line 196 please explain why four times the expected median is the recommendation of the authors.

We have made the particular assumption that since the data reported here evaluate as normally distributed with regard to time, size, and BMI then the medians and the means will be equal. 99.7% of a normal distribution will be within 3 standard deviations. In a standardized normal distribution, the mean is zero and the standard deviation is 1, so that a multiplier greater than 3 should include virtually all of the distribution. By rounding to 4, we extrapolate that all of the distribution of resolving cysts can be covered. We have added this explanation to lines 226-238. After how many years did ultrasounds stop? There was not a stopping point for ultrasounds. What is the percentage of patients that say resolution after 4 years or 6 years? For the largest cysts, 60% would resolve in 4 years and 75% in 6 years. Across all cysts, 80-85% would resolve in 4 years and 90-95% in 6 years.

- Line 216 do the authors believe that the statistical significant findings in some categories such as HRT use or cyst type merit an alternate time course for follow up imaging. Yes, we do. We have presented a data-driven approach which contrasts views by expert ultrasonographers that are weighted to the cases they see the most frequently.

As stated in our clinical implication section (Lines 247-257), our paper provides information to the practicing physician for reference when planning follow-up imaging and counseling depending on the characteristics of an individual’s cyst and medical history. We note that counseling and expectations can be adjusted for age, cyst type, and HRT use.

Overall, this a thought provoking paper that highlights some of the limitations of ovarian cyst follow up. Particularly in premenopausal women this is challenging since
cyst formation is multifactorial. I believe the readers of this journal appreciate this perspective but wonder about the clinical applicability of some of the presented data and wonder if the authors believe any follow up ultrasounds for simple cysts required. **We strongly believe that follow up should be routine.**

*The current guidelines from ACOG recommends repeat imaging for cysts up to 10cm, but the time interval for follow-up has not been established (Lines 221-224). We believe that follow-up ultrasounds should in practice be for simple cysts, however, the characteristics and individual’s medical history should be used as guidance when the practicing physician is deciding on a time interval (Lines 221-224).*
Reviewer #2:

1. Precis - include all seven factors analyzed

_The precis now reads:_

To estimate surveillance times of simple ovarian cysts, resolution times were determined as a function of cyst diameter, cyst structure, age, body habitus, hormone replacement therapy, familial cancer history, and menopausal status.

2. Lines 25-26 states the following: "Individuals who reported a family history of ovarian cancer (64.3%) were more likely to resolve than those that did not (35.7%) (p<0.05). Menopausal status, BMI, or family history made no difference on resolution time (p>0.05). Is family history predictive?

_This is further elaborated in Lines 190-193 which states “Individuals with a family history of ovarian cancer were more likely to have cysts that resolved (64.3% (640/996) vs 35.7% (356/996)); however, resolution times of those with a family history were not significantly different from those without a family history (p<0.05, Table 3).”_

_Meaning, that even though those who had a family history of ovarian cancer were more likely to resolve than those who did not, the times to resolution of these individuals were not significantly different so that family history was not predictive of resolution time._

3. How did you confirm family history of ovarian cancer and hormonal replacement therapy?

_This information was self-reported, however every attempt was made to have the pathology report of their relative submitted to the program. Lines 73-75 states “All study participants completed a questionnaire that included medical history, surgical history, menopausal status, hormone use, and family history of cancer, as previously published (13).”_

4. Line 68 - How did you determine menopausal status for patients s/p hysterectomy?
Postmenopausal status was determined by a self-reporting questioner that asked individuals when they last experienced menses. If it had been greater than 12 consecutive months, then those individuals were classified as postmenopausal. This definition is described by the National Institute of Aging, suited for self-reporting, inclusive, and regardless of their previous surgical history. See lines 75-78

Lines 111-113 states, “Only individuals that had both ovaries visible on their initial TVUS encounter were included in this study and followed over the course of subsequent TVUS examinations”. Lines 115-117 states that “If individuals underwent pelvic surgery, they were censored from the present analysis so that surgical interventions did not influence estimates of cyst resolution.” Lines 113-115 have been added to respond to the reviewer’s query, “Individuals were eligible to enroll in the screening protocol if they had received a hysterectomy at least 12 months prior to their first ultrasound screen.”

5. Line 71 - How many ultrasound technicians performed scans and were they all trained to perform scans using methodology described?

“36 ultrasonographers performed the TVUS over the course of 31 years of the screening program operation, All ultrasonographers that participated in this study were ARDMS certified. All were trained at the University of Kentucky to perform scans using the methodology described.” This statement has been added at lines 105-108.


Lines 101-105 now reads: A normal ovary is ovoid in shape and will have an echotexture with a central echogenic medulla (39). Changes in echogenicity will involve hypoechogenic cystic fluid, and hyperechogenicity for septations within cysts, papillary projections and solid masses (39). In general, normal ovarian size for premenopausal women was considered to be < 20 cm³ and <10 cm³ for postmenopausal women.
7. Line 196 - How did you conclude surveillance should continue four times beyond the expected median resolution time?

*We have made the particular assumption that since the data reported here evaluate as normally distributed with regard to time, size, and BMI, then the medians and the means will be similar. 99.7% of a normal distribution will be within 3 standard deviations. In a standardized normal distribution, the mean is zero and the standard deviation is 1, so that a multiplier greater than 3 should include virtually all of the distribution. By rounding to 4, we extrapolate that all of the distribution of resolving cysts can be covered. We have added this explanation to lines 226-238. Using the median values with a multiplier equal to 1, it can be expected that 50% resolution will occur.*
Reviewer #3:
Review of Manuscript ONG-23-928 "Predicting the resolution and persistence of ovarian cysts: Findings for the practicing physician"

A manuscript that utilizes data on TVUS results gathered from the University of Kentucky Ovarian Cancer screening study has been submitted. As noted by the authors, the objective of the study was to evaluate resolution of incident cysts that occurred during prospective surveillance. The authors subsequently performed analyses to determine the potential factors that may predict resolution as well as timing of said resolution.

They have not included a STROBE checklist which would be expected with the reporting of cohort trial data.

*The Strobe checklist has been included.*

A minor point is that in the manuscript p values are listed both as p <.05 and p < 0.05, please revise to be consistent. I have the following questions and comments.

*The p values have been changed to all reflect p<0.05 or p>0.05*

Précis - This seems to be an incomplete thought. Please revise.

*The precis now states, “To estimate surveillance times of simple ovarian cysts, resolution times were determined as a function of cyst diameter, cyst structure, age, body habitus, hormone replacement therapy, familial cancer history, and menopausal status.”*

Abstract - If space allows add the years to the abstract. Also note in the abstract these are results from a prospective cohort study.

*This has been added in lines 11-13: “47,762 individuals enrolled over 30 years in the prospective cohort University of Kentucky Ovarian Cancer Screening Trial identified 2638 individuals with incident cysts.”*
Line 12 - minor issue but "t" for t-test should be in italics I believe.

*The “t” for t-test has been formatted to italics in lines 133.*

Introduction - Line 50 - minor point but concern for malignancy may be present even if less than 10cm so perhaps remove the size limitation from this comment and/or split this into 2 reasons for surgery - concern for cancer and size > 10cm.

*We agree with the reviewer. Line 55-57 now reads:*

*"Characteristics that are highly suspicious of malignancy include increasing size, the inclusion of papillary or solid components, mass irregularity, presence of ascites, and high internal color Doppler flow"*

Methods - Line 62-6 - What about prior Gyn surgery? Were women with a hysterectomy eligible to enroll? How about a prior USO with no prior surgeries?

*Individuals were eligible to enroll in the screening study if they had received a hysterectomy at least 12 months prior to their first ultrasound screen. However, they were censored from the current analysis in the even that events related to surgery like pelvic adhesions might influence cyst resolution.*

*Lines 111-113 states, “Only individuals with two visible ovaries on their first TVUS encounter were utilized in this study and then followed over the course of subsequent examinations by TVUS”, therefore patients who had one ovary were not eligible. The additional test has been added at lines 113-115, “Individuals were eligible to enroll in the screening protocol if they had received a hysterectomy at least 12 months prior to their first ultrasound screen. Furthermore, lines 115-117 states, “Individuals who underwent pelvic surgery were censored from the present analysis so that surgical interventions did not influence estimates of resolution”.*
Line 65 - Is this individual BRCA status or family status in terms of a pathogenic variant?

*Individual BRCA status. At line 72 we have clarified this: “personal BRCA status”.*

Line 105 - See previous comments on t-test.

*The “t” for t-test has been formatted to italics in lines 133.*

Results - Line 115 - please state that those with prevalent cysts were excluded from this analysis.

*Lines 143-144 now states “and were excluded from analysis”*

Line 134 - Can you provide a range or IQR for the resolution?

*We have inserted the IQRs in Table 3.*

How often was surgery contemplated for these patients? *All enrollees have their screening results referred to their personal physician. We do not know about a surgical decision until after surgery has been performed. We encourage women with cysts to allow us to perform serial screening rather than have surgery so we do not know how often surgery is “contemplated,” but we have follow-up on how often it is done.*

Is there information on women that could serve as a different "control" that had persistence of benign cysts but that underwent surgery instead of continued TVUS assessments?

*No, the problem of surgery at outside hospitals confounds gathering this information, although we do have the numbers.*

Discussion - Line 195 - So to be clear for your data with resolution of 14+ months the
recommendation for continued surveillance would be 56+ months? The answer that we present in the revision is: 42-56 months (see lines 226-238)

We have made the particular assumption that since the data reported here evaluate as normally distributed with regard to time, size, and BMI then the medians and the means will be similar. 99.7% of a normal distribution will be within 3 standard deviations. In a standardized normal distribution, the mean is zero and the standard deviation is 1, so that a multiplier greater than 3 should include virtually all of the distribution. By rounding to 4, we extrapolate that all of the distribution of resolving cysts can be covered. We have added this explanation to lines 226-238.

Line 201 - Did HRT predate the presence of a cyst?

Yes, HRT always predated the detection of a cyst. We have added text that expands the concept that the reviewer suggests. Lines 241-244 now read “It is possible that the use of HRT at the time of cyst diagnosis may result in a more rapid resolution of ovarian cysts by influencing hormone balance and may provide an alternate treatment solution for individuals who experience persisting cysts.”

Line 204 - So if HRT may promote resolution, why did HRT not prevent occurrence?

There is a partial reduction in cyst occurrence in individuals on HRT (see line 177), so that while complete reduction does not occur, there is some reduction.

The tables and figure are duplicated

Figure 2 & 3 are Kaplan Meier curves for cyst resolution and Tables 1 & 2 are characterizations of participants. Table 3 expresses the univariate analysis that results in terms of median resolution times. There is not a duplication.

Tables - Consider making a supplemental table that provides summary information about these selected patients, what made them eligible, and baseline demographics.
Table 1 has been added and displays the patient demographics, as the number of subjects in each group (HRT use, age, menopausal status, BMI, family history of ovarian cancer).

Table 1 - It seems like much of this information is included in the manuscript. If so, perhaps supplemental.
Tables 2 and 3 - No comments

Table 1 is now Table 2, and other reviewers requested that it should be in the manuscript. Although some of the information is included in the manuscript, not all information of the Table 2 is included in writing. Additionally, we feel that the inclusion of the table in the manuscript rather than the supplemental material is necessary for the expansion of ideas written and makes for easier comprehension.

Figures - Figure 1 - Any information on how many of the 971 with persistent cysts had surgery? At the time of this compilation, none of the 971 individuals reported receiving pelvic surgery
Figures 2 and 3 - No comments
STATISTICAL EDITOR COMMENTS:
lines 22-23 and Table 2: In Table 2, the stats test by age strata was a test for an overall difference, not a specific comparison one age group. Need to provide CIs for each of the age groups (see below comments re: Table 2) and demonstrate whether the p < 0.05 difference was for age 70+ vs each of the other groups, or simply vs the < 40 age group. **The CI have been added to Table 2.**

The CI for <40 is 1.514-2.086; 40-69: 1.180-1.320; 70+ 0.965-1.035. The p<0.05 is significant for the overall ages compared to survival times. The <40 group was not statistically significant regarding time to resolution with a p = 0.146. The other two groups were significant with the group 40-69: p = 0.036 and 70+: p = 0.018.

<table>
<thead>
<tr>
<th>Age</th>
<th>Mean</th>
<th>95% Confidence Interval</th>
<th>Median</th>
<th>95% Confidence Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Lower Bound</td>
<td>Upper Bound</td>
<td></td>
</tr>
<tr>
<td>1.00</td>
<td>2.837</td>
<td>2.462</td>
<td>3.212</td>
<td>1.80</td>
</tr>
<tr>
<td>2.00</td>
<td>2.287</td>
<td>2.174</td>
<td>2.399</td>
<td>1.25</td>
</tr>
<tr>
<td>3.00</td>
<td>1.801</td>
<td>1.472</td>
<td>2.129</td>
<td>1.00</td>
</tr>
<tr>
<td>Overall</td>
<td>2.300</td>
<td>2.198</td>
<td>2.403</td>
<td>1.25</td>
</tr>
</tbody>
</table>

The following has been added to the legend for Figure 2. **“Results were obtained using the Kaplan-Meier method. Numbers at risk for each group at the indicated time points are shown below the x-axis.”**

Fig 2: This appears to be a K-M graph, but need to explicitly label it. Also, need to show the N at risk for each group at the indicated time points along the x-axis. As long as the censoring rates remain similar for the two groups vs time, then the estimates for % remaining or its converse % resolution is valid. If there is a differential rate of censoring, then the estimates are likely biased, since one is assuming an equivalence of outcomes for those shown on the K-M graphs vs those censored from the graphs.

The following has been added to the legend for Figure 2. **“Results were obtained using the Kaplan-Meier method. Numbers at risk for each group at the indicated time points are shown below the x-axis.”**

Fig 3: Again, these appear to be K-M graphs, but they need to be labelled explicitly. Also, need to include either below the xx-axes at the indicated time points, or in a separate Table, the N for those remaining at risk for each group. The following has been added to the legend for Figure 3. **“Results were obtained using the Kaplan-Meier method. Numbers at risk for each group at the indicated time points are shown below the x-axis.”**


Table 1 (now Table 2): Need to include CIs for each of the % resolved. **The CIs have been added.**

Table 2 (now Table 3): Need to include IQRs for each of the median resolution times cited. Should also indicate the N for each subset. **The IQRs have been added as well as N to Table 3.**

Table 3 (now Table 4): Should round all HRs and their CIs to 0.01 precision. **Rounding as recommended has been performed.**

Need to identify the referent group. **The referent groups have been identified in the legend.**

Need to state whether the assumption of proportion hazard rates was examined and verified. **At lines 287-294, the following text has been added that addresses proportional hazards:**

> Certain realizations about the multivariate analysis reported on here are important. The Cox models fitted in Table 4 assume that a proportional hazards assumption is met by each risk factor listed in column one of the table. However, only BMI met this assumption when it was assessed by the supremum test for proportional hazards. All the other risk factors did not meet this assumption in this well known test. Further re-analysis showed that for each risk factor the assumption failed to be met whenever the time to cyst resolution was less than 2.3 years, while for longer resolution times the hazard ratios listed in Table 4 do hold.”

Need units for BMI. **Units for BMI have been added.**

Need to include N for each subset. **N has been included for each subset.**

Need to list the variables retained in the final multivariate models in footnote to Table. **This statement has been added to the table legend:** “All the factors listed were the variables retained in the final multivariate models.”