

Comparing the Diagnostic Accuracy of Symptom Index and the Risk of Malignancy Index with Histopathological Outcomes in Adnexal Masses

Amita Ray

ABSTRACT

Aim: Accurate identification of the nature of an adnexal mass can ensure optimum management and with this aim in mind we applied tests of diagnostic accuracy to the symptom index (SI) and the risk of malignancy index 1 (RMI-1) separately and in combination to differentiate between benign and malignant adnexal masses in women admitted for surgery at three different medical college institutions in India. This was done to frame a referral policy for women with adnexal masses, facilitate triaging and for counseling the woman and her family.

Methods: All the participants were given a survey to complete based on which the participants were either SI positive or negative. Ultrasonographic parameters and CA-125 levels were used to calculate the RMI for each participant. Using Histopathology as the gold standard we compared the sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), positive and negative likelihood ratios and diagnostic accuracy of the SI and RMI separately and in combination.

Results: When compared to the SI the RMI rates higher in all tests of diagnostic accuracy. Sensitivity (SI—41.4% RMI—77.8%) Specificity (SI—73.24%, RMI—80.6%), PPV (SI—38.71% RMI—60.0%), NPV (SI—75.36% RMI—90.6%) and diagnostic accuracy (SI—64% RMI—79.8%) diagnostic accuracy does not improve when the tools are combined.

Conclusion: The RMI can be used as a triaging tool and also for framing a referral policy for adnexal masses. For counseling patients with adnexal masses, the RMI through a better option than the SI does not have a good positive or negative likelihood ratio to either rule out or rule in a diagnosis of malignancy in individual cases.

Clinical significance: Menopausal status, CA-125, and ultrasound features are much more predictive of the nature of an adnexal mass when compared to symptoms making RMI a better tool for triaging and referral and counseling women with adnexal masses

Keywords: Adnexal mass, Risk of malignancy index 1, Symptom index, Tests of diagnostic accuracy.

How to cite this article: Ray A. Comparing the Diagnostic Accuracy of Symptom Index and the Risk of Malignancy Index

with Histopathological Outcomes in Adnexal Masses. *J South Asian Feder Menopause Soc* 2018;6(2):91-96.

Source of support: Nil

Conflict of interest: None

Date of received: 10 January 2018

Date of acceptance: 27 January 2018

Date of publication: July 2018

INTRODUCTION

Ovarian masses are one of the most common problems faced by a practicing gynecologist. They have varied etiologies and knowing the nature of the mass goes a long way in the preparedness for dealing with it. Whether they should be operated, who should operate, where and how the operation needs to be done and how the patient and her attendants are to be counseled are important questions that need answers.¹ Such masses many of which are discovered incidentally, and particularly in premenopausal women need only reassurance and reassessment.

On the contrary, several such masses require surgery which in cases of malignancy may be extensive and associated with significant morbidity.¹⁻³

In case the decision for surgery is made, knowing the nature of the adnexal mass is of utmost importance for the clinician to optimize management. Accurate identification of malignant tumors can lead to the timely referral of patients to oncology centers where a team consisting of the gynecological oncologist, the radiation and medical oncologist can ensure optimum treatment. Such management can positively influence the prognosis. On the other hand, benign masses could be managed by either expectant management or by conservative surgery leading to reduced morbidity and preservation of fertility.¹⁻³ Tools that predict the nature of the mass with reasonable accuracy prove invaluable in improving patient management. Various combined methods for evaluating the nature of and the risk of malignancy in an adnexal mass have been developed. The variables included are clinical symptoms as in the SI,⁴ or a combination of menopausal status, sonographic features and CA-125 levels as in the RMI.⁵

In this study, we evaluated and compared the ability of the SI and RMI1 separately and in combination to differ-

Professor and Head

Department of Obstetrics and Gynecology, IQ City Medical College, Durgapur, West Bengal, India

Corresponding Author: Amita Ray, Professor and Head, Department of Obstetrics and Gynecology, IQ City Medical College, Durgapur, West Bengal, India, e-mail: amitarays@gmail.com,

entiate between benign and malignant adnexal masses in patients from three different medical college institutions in India. This was done to help frame a referral policy for women with adnexal masses, facilitate triaging of women with such masses and counseling her and her family

METHODS

Data Collection

Women with adnexal masses admitted to the three affiliated hospitals for surgical management were included in this study. Women with adnexal masses who were either managed conservatively were not fit for surgery or denied surgery was excluded. After informed consent (Appendix 1) all of the participants were given a survey to complete. The survey asked for age, the last menstrual period and the eight symptoms included in the modified Goff’s symptom Index (Appendix 2). Based on the survey women were either SI positive or negative. Transvaginal ultrasound was performed and if the mass was too big to be observed completely a transabdominal scan was also done serum CA-125 levels were measured using electro-chemi-chemiluminescent immuno assay (ECLIA).

Based on the data collected the RMI-1 Score for each participant was calculated (Appendix 3).

Data Analysis

Several measures relate to the different aspects of the diagnostic procedure. Discriminatory tests like sensitivity and specificity are not affected by disease prevalence whereas tests used to assess the predictive ability are affected by disease prevalence.^{6,7} For our study, we calculated the sensitivity, specificity, positive and negative likelihood ratios, positive predictive value (PPV) and negative predictive value (NPV) and diagnostic accuracy. Sensitivity, specificity and the likelihood ratios are discriminatory methods and can be used to frame policies for triaging patients. Since these values are not affected by disease prevalence, they can also be extrapolated to other populations.^{6,7} While PPV, NPV and diagnostic accuracy are predictive methods which are useful in predicting the probability of malignancy in an individual case and could help the clinician in counseling the woman and her family. Since these values are dependent on disease prevalence, they are useful in knowing the probability of malignancy in a particular individual and are not applicable to other settings.

In this study of diagnostic accuracy, we have included all the items listed in standards for reporting diagnostic accuracy studies (STARD) (Appendix 4). This list of items was developed to contribute to the completeness and transparency of diagnostic accuracy studies.^{8,9}

RESULTS

The mean age of the participants was 44.65 +/-14.86 years (Table 1).

Using histopathology as the gold standard for diagnosing malignancy the SI had (a) 12 true positives, (b) 19 false positives, (c) 17 true negatives, and (d) 52 false negatives (Table 2).

The SI demonstrated a sensitivity of 41.38% ; specificity of 73.24%, PPV 38.71%, NPV 75.36%, positive likelihood ratio of 1.55 , a negative likelihood ratio of 0.80 and an accuracy of 64.00% in our study population (Table 3)

Using histopathology as the gold standard for diagnosing malignancy the RMI had (a) 21 true positives, (b) 14 false positives, (c) 6 true negatives, and (d) 58 false negatives (Table 4).

The RMI-1 model demonstrated a sensitivity of 77.8%, specificity of 80.6%, PPV 60%, NPV 90.6%, positive likelihood ratio of 4, a negative likelihood ratio of 2.5 and an accuracy of 79.8% in our study population (Table 5).

Table 1: Age distribution

	Mean	Std. deviation
Age	44.65	14.86

Table 2: Diagnosis of malignancy by symptom index (SI) and histopathological examination (HPE)

		HPE		Total
		Cancer	Not cancer	
Symptom index (SI)	Cancer	12 (a)	19 (b)	31
	No cancer	17 (c)	52 (d)	69
	Total	29	71	100

Table 3: Values for tests of diagnostic accuracy for symptom index (SI)

	Formula	Value	95% CI
Sensitivity	$\frac{a}{a + b}$	41.38%	23.52 to 61.06%
Specificity	$\frac{d}{c + d}$	73.24%	61.41 to 83.06%
Positive likelihood ratio	$\frac{\text{Sensitivity}}{1 - \text{Specificity}}$	1.55	0.87 to 2.76
Negative likelihood ratio	$\frac{1 - \text{Sensitivity}}{\text{Specificity}}$	0.80	0.57 to 1.12
Disease prevalence	$\frac{a + b}{a + b + c + d}$	29.00%	20.36 to 38.93%
Positive predictive value	$\frac{a}{a + c}$	38.71%	26.13 to 52.99%
Negative predictive value	$\frac{d}{b + d}$	75.36%	68.60 to 81.07%
Accuracy	$\frac{a + d}{a + b + c + d}$	64.00%	53.79 to 73.36%

Comparing the Diagnostic Accuracy of Symptom Index and the Risk of Malignancy Index

Table 4: Diagnosis of malignancy by risk of malignancy index (RMI) and histopathological examination (HPE)

		HPE		Total
		Cancer	Not cancer	
RMI	Cancer	21(a)	14(b)	35
	No Cancer	6 (c)	58 (d)	65
	Total	27	73	100

Table 6: Diagnosis of malignancy by a combination of symptom index (SI) and risk of malignancy index (RMI) vs. histopathological examination (HPE)

For the combination to be positive—The woman has to be symptom index positive and have an RMI above 200

		HPE		Total
		Cancer	Not cancer	
Symptom index + and RMI above 200	Cancer	12 (a)	6 (b)	18
	Not cancer	17 (c)	67 (d)	82
	Total	29	71	100

Using histopathology as the gold standard for diagnosing malignancy the combination had (a) 12 true positives, (b) 6 false positives, (c) 17 true negatives, and (d) 67 false negatives (Table 6).

The combination demonstrated a sensitivity of 41.38%, specificity of 91.78%, PPV 66.67%, NPV 79.76%, positive likelihood ratio of 5.03, a negative likelihood ratio of 0.64 and an accuracy of 77.45% in our study population (Table 7).

DISCUSSION

Overall this study has revealed that RMI by itself is more useful when compared to the SI and a combination of SI and RMI in assessing the nature of a pelvic mass.

The SI had a low sensitivity (41.38%) and a moderate specificity (73.24%). The positive and the negative predictive values of SI are 38.7% and 75.3%, respectively. Taken together these diagnostic parameters convey to us that symptoms of ovarian malignancy (based on which the SI has been designed) are not pathognomonic of the disease and can be mimicked by other conditions (higher false positive and thus lesser specificity). In our study, many participants with endometriosis had a positive symptom index. At the same time often in the presence of malignancy these symptoms are absent (higher false negative and thus lesser sensitivity) and because of which the ovarian malignancy has been called a silent killer.

The RMI gave a higher sensitivity (70.8%) and specificity (80.1%) as well as a better positive predictive (60%) and negative predictive value (90%). The results convey that the RMI is a better tool when compared to SI.⁷ Similar figures for the RMI-1 has been reported by other studies an indication that sensitivity and specificity are not affected by disease prevalence.¹⁰⁻¹³

Table 5: Values of tests of diagnostic accuracy for the Risk of malignancy index (RMI)

	Formula	Value	95% CI
Sensitivity	$\frac{a}{a + b}$	77.78%	57.74 to 91.38%
Specificity	$\frac{d}{c + d}$	80.56%	69.53 to 88.94%
Positive likelihood ratio	$\frac{Sensitivity}{1 - Specificity}$	4.00	2.40 to 6.67
Negative likelihood ratio	$\frac{1 - Sensitivity}{Specificity}$	0.28	0.13 to 0.56
Disease prevalence	$\frac{a + b}{a + b + c + d}$	27.27%	18.80 to 37.15%
Positive predictive value	$\frac{a}{a + c}$	60.00%	47.35 to 71.44%
Negative predictive value	$\frac{d}{b + d}$	90.63%	82.55 to 95.18%
Accuracy	$\frac{a + d}{a + b + c + d}$	79.80%	70.54 to 87.20%

Table 7: Values of tests of diagnostic accuracy for a combination of symptom index (SI) and risk of malignancy index (RMI)

	Formula	Value	95% CI
Sensitivity	$\frac{a}{a + b}$	41.38%	23.52% to 61.06%
Specificity	$\frac{d}{c + d}$	91.78%	82.96 to 96.92%
Positive likelihood ratio	$\frac{Sensitivity}{1 - Specificity}$	5.03	2.09 to 12.14
Negative likelihood ratio	$\frac{1 - Sensitivity}{Specificity}$	0.64	0.47 to 0.87
Disease prevalence	$\frac{a + b}{a + b + c + d}$	28.43%	19.94 to 38.22%
Positive Predictive Value	$\frac{a}{a + c}$	66.67%	45.33% to 82.83%
Negative predictive value	$\frac{d}{b + d}$	79.76%	74.23 to 84.35%
Diagnostic accuracy	$\frac{a + d}{a + b + c + d}$	77.45%	68.11 to 85.14%

As has been noted in other studies the sensitivity of RMI is less because many germ cell malignancies do not have an elevated CA-125 and are not in the menopausal age group which gives rise to false negatives and although the specificity of RMI is high endometriosis which can both raise the CA-125 and gives a multiloculated appearance in an adnexal mass could give rise to false positives.^{11,12}

In our study we found the RMI to have a high negative predictive value meaning that if the RMI falls below 200 the probability that the person does not have cancer is high (90.3%). By comparison, the SI has a 75% nega-

tive predictive value. Thus an RMI of less than 200 gives a better probability of a benign lesion than a negative SI, and this can be used when counseling women and their family. The positive predictive values of both the tools were low and thus cannot be used with confidence to predict the probability of malignancy in an adnexal mass. Both positive and negative values are dependent on disease prevalence, and our findings of this study as regards negative and positive predictive values cannot be extrapolated to other populations and therefore we have not compared these findings with other such studies.^{6,7}

The positive likelihood ratio (which is the best method of confirming the diagnosis of malignancy) is greater for the RMI than SI. But the value is 4.00 meaning that an RMI above 200 does not confirm malignancy. A positive likelihood ratio of 10 would have been almost confirmatory of a diagnosis of malignancy.^{6,7}

The negative likelihood ratio (which is the best method of ruling out a diagnosis of malignancy) is lesser for the RMI as compared to SI. But again, the value is not meaningful, and so an RMI below 200 does not rule out malignancy. A negative likelihood ratio of 0.1 would be confirmatory of ruling out the diagnosis of malignancy.^{6,7}

Likelihood ratios are not affected by disease prevalence, and so these findings could be extrapolated to other settings. Our literature search did not reveal any study which has compared the positive and negative likelihood ratios of either SI or RMI.^{6,7}

We combined the results of SI and RMI to see whether the tests of diagnostic accuracy improve. A woman who was symptom Index positive and had an RMI of above 200 was taken as having a diagnosis of malignancy. We justify this combination because none of the variables in these two tools are common combining, however, did not improve the tests of diagnostic accuracy

CONCLUSION

The RMI as a diagnostic tool to assess the nature of an adnexal mass outperforms the SI on almost all tests of diagnostic accuracy. Thus, RMI can be used as a tool for triaging patients and frame referral policies of a hospital. A center which does not have oncology, radiation and chemotherapy facilities should refer women with adnexal masses having an RMI of greater than 200 to such centers

that have them. The risk of the mass being malignant far outweighs the chance of it being benign and optimal debulking followed by chemotherapy will improve the overall survival (OS).

When it comes to counseling individual patients since the RMI has a high NPV an RMI value of less than 200 can be definitely used to reassure the patient and her relatives. But the same cannot be said for the positive predictive value

REFERENCES

1. Gillis CR, Hole DJ, et al. Medical audit, cancer registration, and survival in ovarian cancer. *Lancet* 1991;337:611-612.
2. Eisenkop SM, Spirtos NM, et al. The impact of subspecialty training on the management of advanced ovarian cancer. *Gynecol Oncol* 1992;47:203-209.
3. Kehoe S, Powell J, et al. The influence of the operating surgeon's specialisation on patient survival in ovarian carcinoma. *Br J Cancer* 1994;70:1014-1017.
4. Goff BA, Mandel LS, et al. Development of an ovarian cancer symptom index: Possibilities for earlier detection. *Cancer* 2007;109:221-227.
5. Jacobs I, Oram D, et al. A risk of malignancy index incorporating CA-125, ultrasound and menopausal status for the accurate preoperative diagnosis of ovarian cancer. *Br J Obstet Gynaecol* 1990;97:922-929.
6. Irwig L, Bossuyt P, et al. Designing studies to ensure that estimates of test accuracy are transferable. *BMJ* 2002;324(7338):669-671.
7. Raslich MA, Markert RJ, et al. Selecting and interpreting diagnostic tests. *Biochemia Medica* 2007;17(2):139-270.
8. Bossuyt PM, Reitsma JB, et al. The STARD statement for reporting studies of diagnostic accuracy: explanation and elaboration. *Clin Chem* 2003;49:7-18.
9. Wilczynski NL. Quality of reporting of diagnostic accuracy studies: no change since STARD statement publication-before-and-after study. *Radiology* 2008;248(3):817-823.
10. Obeidat B, Amarin Z, et al. Risk of malignancy index in the preoperative evaluation of pelvic masses. *International Journal of Gynecology & Obstetrics* 2004;85(3):255-258.
11. Yamamoto Y, Yamada R, et al. Comparison of four malignancy risk indices in the preoperative evaluation of patients with pelvic masses. *European Journal of Obstetrics & Gynecology and Reproductive Biology* 2009;144(2):163-167.
12. van den Akker PA, Aalders AL, et al. Evaluation of the risk of malignancy index in daily clinical management of adnexal masses. *Gynecol Oncol* 2010;116(3):384-388.
13. Manjunath A, Sujatha K, et al. Comparison of three risk of malignancy indices in evaluation of pelvic masses. *Gynecol Oncol* 2001;81(2):225-229.

Appendix 1**Consent Form****Dear Participants**

This is a study on ovarian tumors being done at three medical colleges. As per this study all women like you who have an ovarian tumor and are about to undergo surgery will be required to fill a questionnaire.

The questionnaire will ask you to answer certain questions related to the problems or symptoms that you are experiencing over the past one year. The questionnaire does not ask you to identify yourself in any way.

The information provided by you in this questionnaire will be used for research purposes. It will not be used in a manner which would allow identification of your individual responses.

Yours,

Dr Amita Ray

Principal Researcher

e-mail: dramitarays@rediffmail.com

Phone: 9652068021

I give my consent for participation in this educational project. I understand that the data collected from the questionnaire will be used for research purposes

Appendix 2**Modified Goff's Symptom Index**

An adnexal mass was considered to be malignant, if any of the following symptoms occurred for more than 12 times in a month for a period of less than 1 year:

- Pelvic pain
- Abdominal pain,
- Increased abdominal size
- Bloating,
- Difficulty eating
- Feeling full.
- Urinary urgency
- Increased frequency.

Appendix 3

Risk of malignancy index 1 (Jacobs et al. 1990)

The risk of malignancy index was calculated by:

$U \times M \times CA125$, where a total ultrasound score of 0 made $U = 0$, a score of 1 made $U = 1$, and a score of ≥ 2 made $U = 3$; premenopausal status made $M = 1$ and postmenopausal $M = 3$. The serum level of CA125 was applied directly to the calculation. The cut off level was selected at 200 above which was considered to indicate malignancy.

Appendix 4: STARD Guidelines

<i>Our study</i>	<i>STARD guidelines</i>
<i>Tests of diagnostic accuracy</i>	
1. Sensitivity, specificity, likelihood ratios, positive and negative pred	Identification as a study of diagnostic accuracy using at least one measure of accuracy (such as sensitivity, specificity, predictive values, or AUC)
<i>Abstract</i>	
2. Abstract given as per STARD guidelines	Structured summary of study design, methods, results, and conclusions (for specific guidance, see <i>STARD for Abstracts</i>)
<i>Introduction</i>	
3. Introduction as per STARD guidelines	Scientific and clinical background, including the intended use and clinical role of the index test
4. To compare the ability of the SI and RMI 1 separately and in combination to differentiate between benign and malignant adnexal masses	Study objectives and hypotheses
<i>Methods</i>	
<i>5. Study design</i>	
Data collection planned before index test (prospective study)	Whether data collection was planned before the index test and reference standard were performed (prospective study) or after (retrospective study)
<i>6. Participants</i>	
<i>Eligibility criteria</i>	
7. All women with a diagnosis of an adnexal mass scheduled for surgery	On what basis potentially eligible participants were identified (such as symptoms, results from previous tests, inclusion in registry)
8. Three centers from 2017–2018	Where and when potentially eligible participants were identified (setting, location and dates)
9. Consecutive	Whether participants formed a consecutive, random or convenience series
<i>Test methods</i>	
10a Symptom Index SI (detailed in Appendix) risk of malignancy index RMI (detailed in Appendix)	Index test, in sufficient detail to allow replication
10b Reference standard histopathology	Reference standard, in sufficient detail to allow replication
11. Histopathology golden and final diagnosis alternatives do not exist	Rationale for choosing the reference standard (if alternatives exist)
12a Cut off for RMI as specified by RCOG is 200	Definition of and rationale for test positivity cut-offs or result categories of the index test
The symptom index is either positive or negative and does not have a cutoff reference cited	
12b	Definition of and rationale for test positivity cutoffs or result categories of the reference standard, distinguishing pre-specified from exploratory
No cut offs in Histopathology it would be either yes for malignancy and no for benign	
13a	Whether clinical information and reference standard results were available to the performers/readers of the index test
No	
13b	Whether clinical information and index test results were available to the assessors of the reference standard
No	
<i>Analysis</i>	
14. Statistical software for calculation of all the tests specified above	Methods for estimating or comparing measures of diagnostic accuracy
15. There were no indeterminate tests	How indeterminate index test or reference standard results were handled
16. There were no missing data	How missing data on the index test and reference standard were handled
17. No variability	Any analyses of variability in diagnostic accuracy, distinguishing pre-specified from exploratory
18	Intended sample size and how it was determined
<i>Results</i>	
<i>Participants</i>	
20. Specified in results section	Baseline demographic and clinical characteristics of participants
21a Not assessed	Distribution of severity of disease in those with the target condition
21b Endometriosis	Distribution of alternative diagnoses in those without the target condition
22. Time interval mean 7 days no clinical interventions in between	Time interval and any clinical interventions between index test and reference standard
<i>Test results</i>	
23. Specified in results section	Cross tabulation of the index test results (or their distribution) by the results of the reference standard
24. Specified in results section	Estimates of diagnostic accuracy and their precision (such as 95% confidence intervals)
25. Nil	Any adverse events from performing the index test or the reference standard
<i>Discussion</i>	
26	Study limitations, including sources of potential bias, statistical uncertainty, and generalisability
Specified in discussion	
27	Implications for practice, including the intended use and clinical role of the index test
Specified in discussion	
<i>Other Information</i>	
No sources of funding	