Concordance Between İntraoperative Frozen Section And Final Pathology İn Adnexal Masses Using The Kappa Test

Adneksiyal Kitlelerde, İntraoperatif Frozen Section İle Final Patoloji Arasındaki Uyumda Kappa Testinin Kullanımı

Hakan Cokmez[1], Çetin Aydın[1], İrfan Öcal[2], Simge Tezel Yozgat[1]

[1]İzmir Atatürk Eğitim Ve Araştırma Hastanesi, Kadın Hastalıkları Ve Doğum Kliniği, İzmir
[2]İzmir Katip Çelebi Üniversitesi Tıp Fakültesi, Patoloji Ana Bilim Dalı, İzmir

Dergiye Ulaşma Tarihi: 10.09.2019 Dergiye Kabul Tarihi: 18.02.2020 Doi: 10.5505/aot.2020.24085

ÖZET

GİRİŞ ve AMAÇ: Adneksiyal kitlelerin teşhisi için intraoperatif frozen section (IFS) ile final histopatolojik sonuçlar arasındaki uyumu değerlendirmek.

YÖNTEM ve GEREÇLER: Adneksiyal kitle nedeniyle IFS incelemesi yapılan 274 hastanın verileri retrospektif olarak toplandı. Adneksiyal kitlelerin teşhisinde IFS ve final histopatolojik sonuçlarının uyumu, rastgele uyum oranını düzelterek hesaplayan kappa testi kullanılarak değerlendirildi.

BULGULAR: IFS ve final patoloji sonuçları arasındaki genel uyum oranı (%94,0), Cohen's kappa tablosu ile değerlendirildiğinde güçlü bir uyumu göstermiştir (K = 0,847; p <0,001). IFS sensitivite değerleri benign, borderline ve malign vakalarda sırasıyla %99,0; %68,4 ve %84,0 idi. Spesifite oranları benign, borderline ve malign vakalarda sırasıyla %84,1; %99,2 ve %98,7 idi. Benign, borderline ve malign vakalarda prevalans oranları sırasıyla %74,8; %6,9 ve %18,3 idi.

TARTIŞMA ve SONUÇ: Adneksiyal kitlelerin teşhisi için IFS ile final patoloji sonuçları arasındaki güçlü uyum, IFS'nin en düşük prevalansa sahip olan borderline vakalar için bile güvenilir olduğunu göstermiştir.

Anahtar Kelimeler: frozen section, histolojik teknikler, ovaryan neoplazmlar

ABSTRAC

INTRODUCTION: To evaluate the concordance between the intraoperative frozen section (IFS) and final histopathological results for diagnosing adnexal masses.

METHODS: The data of 274 patients who underwent IFS examination of an adnexal mass were collected retrospectively. The concordance of IFS and final histopathological results for diagnosing adnexal masses was evaluated using the Kappa test, which was calculated by correcting the random fit rate.

RESULTS: The overall agreement rate (94.0%) between the IFS and final pathology results showed strong agreement when evaluated with Cohen's kappa table (K = 0.847; p < 0.001). The IFS sensitivity values were 99.0%, 68.4%, and 84.0% for benign, borderline, and malignant cases, respectively. Specificity ratios were 84.1%, 99.2%, and 98.7% for benign, borderline, and malignant cases, respectively. Prevalence rates were 74.8%, 6.9%, and 18.3% for benign, borderline, and malignant cases, respectively.

DISCUSSION AND CONCLUSION: Strong concordance between IFS and final pathology results for diagnosing adnexal masses showed that IFS is reliable even for borderline cases, which have the lowest prevalence.

Keywords: frozen section, histological techniques, ovarian neoplasms

INTRODUCTION

All gynecologic surgeons want to rule out ovarian cancer as the differential diagnosis of adnexal masses because its 5-year survival rate is less than 45% and it is the eighth most common cause of cancer death in women (1). Clinical examinations, ultrasound examinations, and analyses of tumor markers are used to evaluate adnexal masses, but none is sensitive or specific enough to detect malignancy (2). Intraoperative frozen section (IFS) is a rapid pathological method that can define the histology and character (benign, borderline, malignant) of adnexal masses during surgery. Intraoperative decision-making

Adress for correspondence: Izmir Ataturk Education and Research Hospital, Karabaglar 35360 Izmir - Türkiye
e-mail: hakancokmez@hotmail.com
Available at www.actaoncologicaturcica.com

during surgical management of adnexal masses, which may vary from mass excision in benign cases to debulking in malignant cases, is usually performed according to the results of the frozen section. To prevent morbidity and mortality due to incomplete surgery or unnecessary extensive surgery, high rates of sensitivity and specificity are necessary and, with the IFS method, expected.

In literature, the role of IFS in the surgical management of adnexal masses has been evaluated, and its positive predictive value (PPV) has been found to vary widely (3,4). PPV is affected by the prevalence (5). For borderline histology specific to epithelial ovarian tumors, which has a lower prevalence than benign and malignant histology, the PPV of IFS is low (6,7). However, because of the low recurrence rates of borderline ovarian tumors, conservative surgery is recommended, especially for patients who wish to preserve fertility (8); therefore, an accurate intraoperative diagnosis of borderline ovarian tumors is important.

This study aimed to evaluate the correlation between the IFS and final pathology results for the histopathological diagnosis of adnexal masses, and to evaluate the Kappa test results, which were corrected according to the random rate and were not prevalence-dependent.

MATERIALS and METHODS

Data were collected retrospectively from patients' medical records at our institution between January 1, 2009 and January 1, 2019; IFS was requested for the evaluation of an adnexal mass during surgery. Before starting our study, approval was obtained from the ethics committee of our institution (#403/2018). The study was performed in accordance with the ethical standards described in an appropriate version of the 1964 Declaration of Helsinki, as revised in 2013. Because of the retrospective design of the study and anonymized data used in the analyses, informed consent was not obtained from the patients.

All frozen section materials were evaluated by pathologists or gynecopathologists working in our clinic to determine the tumor localization, maximum diameter, color, content, heterogeneity, infiltration pattern, and capsule integrity. Depending on the size and heterogeneity of the tumor, one to four sections were sampled in a cryostat; for the pathological examination, one section at least per cm was obtained from the patients in whom the tumor diameter was ≤ 5 cm, in which every 5 cm increased in mass size would increased one more the number of sections per cm (max four section). All sections were stained with hematoxylin and eosin. After the sections were evaluated, the diagnosis based on IFS was communicated to the surgical team by telephone. According to the frozen section report, appropriate surgical procedures were continued. The final pathologic diagnosis was accepted by a gynecopathologist with at least 5 years of experience. Records for all patients in the benign and malignant groups, which were defined according to IFS and final pathology results, included data for all epithelial, sex cord stromal, and germ cell tumors. The only valid data for the borderline group were those of epithelial tumors.

A comparative analysis was performed after the cases were grouped as "agreement" or "no agreement" according to IFS and final pathology results. Agreement between IFS and final pathology was evaluated by the kappa test. The level of agreement indicated by the kappa value was expressed according to the generally accepted Cohen's kappa table (9) (Table 1).

Additionally, the final pathology results for the benign, borderline, and malignant groups were considered accurate, and the following predictive tests were performed using a 2×2 table (5) (Table 2) and expressed as a percentage

- Sensitivity (true positives / true positives + false negatives)
- Specificity (true negatives / true negatives + false positives)
- Accuracy [(true positives + true negatives) / overall]
- PPV (true positives / true positives + false positives)
- Negative predictive value (true negatives / true negatives + false negatives).

Prevalence [(true positives + false negatives) / overall] values were expressed as a percentage. Age, preoperative cancer antigen 125 (CA-125) levels, mass size, histological type, neoplastic status, and character data were recorded. The variables were investigated using visual (histograms, probability plots) and analytical methods (Kolmogorov-Smirnov/Shapiro-

Adress for correspondence: Izmir Ataturk Education and Research Hospital, Karabaglar 35360 Izmir - Türkiye
e-mail: hakancokmez@hotmail.com
Available at www.actaoncologicaturcica.com

Wilk's test) to determine whether they were normally distributed. Because data on age, tumor size, and CA-125 levels were not normally distributed, these parameters were compared among the groups (benign, borderline, and malignant) using the Kruskal-Wallis test ($p < 0.05$). The Mann-Whitney U test was used to determine which groups (benign and malignant) had significant differences. Then, the new p-value was determined by Bonferroni correction (the new p-value for the three [benign, borderline, malignant] groups was determined as follows: 0.05 / 3 = 0.017) to evaluate significant differences between the benign and malignant groups. Statistical calculations were performed using IBM SPSS Statistics for Windows version 20.0 (IBM Corp., Armonk, NY).

Table 1: Cohen's kappa interpretation values

k-value	Agreement level	Reliable data, %
0 – 0.20	None	0 – 4
0.21 – 0.39	Minimal	4 – 15
0.40 – 0.59	Weak	15 – 35
0.60 – 0.79	Moderate	35 – 63
0.80 – 0.90	Strong	64 – 81
>0.90	Almost perfect	82 – 100

Table 2: 2×2 table

	Final pathology		
Frozen section	Data reviewed: Positive	Data reviewed: Negative	Total
Data reviewed: Positive	TP	FP	TP+FP
Data reviewed: Negative	FN	TN	FN+TN
Total	TP+FN	FP+TN	TP+FP+TN+FN

Abbreviations: FN, false negative; FP, false positive; TN, true negative; TP, true positive.

RESULTS

This study included 274 patients who underwent surgery and IFS for an adnexal mass at our institution. The overall agreement between the frozen section and final pathology results was 94.2% ($k = 0.847$; $p < 0.001$). When the k-value was evaluated according to Cohen's kappa table, there was strong agreement. The distribution of IFS and final pathology results among benign, borderline, and malignant groups is shown in Table 3.

Table 4 shows the results of the evaluation of the predictability of IFS according to the final pathology results. Of the 16 (5.8%) cases with no agreement between IFS and final pathology, 12 (4.4%) were underdiagnosed and 4 (1.5%) were overdiagnosed.

The mean age of the patients was 52.0 ± 13.9 years (minimum age, 15 years; maximum age, 83 years). There was a statistically significant difference in age among the benign, borderline, and malignant groups according to IFS ($p = 0.009$) and final pathology ($p = 0.001$) results. This significant difference was due to the differences between the benign and malignant groups according to the IFS and final pathology results ($p = 0.002$ and $p < 0.001$, respectively).

The mean tumor size was 9.8 ± 5.6 cm. The frozen section results indicated no significant difference in the tumor size of the benign, borderline, and malignant groups ($p = 0.129$). The final pathology results indicated a significant difference in the tumor size of the benign, borderline, and malignant groups ($p =$

Adress for correspondence: Izmir Ataturk Education and Research Hospital, Karabaglar 35360 Izmir - Türkiye
e-mail: hakancokmez@hotmail.com
Available at www.actaoncologicaturcica.com

0.02). This significant difference was due to the difference between benign and borderline cases (p = 0.019). However, when Bonferroni correction was used, the difference was found to be larger than that of the new p-value (0.017); therefore, it was statistically insignificant.

The mean CA-125 level was 25.07 ± 24.04 U/mL. The frozen section results indicated no significant relationship between the CA-125 levels and the benign, borderline, and malignant groups (p = 0.184). The final pathology results, however, indicated a significant difference between the CA-125 levels and the benign, borderline, and malignant groups (p = 0.029). This significant difference seemed to be due to the differences between the benign and malignant groups (p = 0.021). However, when the Bonferroni correction was used, the difference was greater than the new p-value (0.017); therefore, it was statistically insignificant. Age, tumor size, and distribution of CA-125 levels of the groups are shown in Figure 1.

When the agreement status of the IFS and final pathology results was compared for age, tumor size, and CA-125 levels (Table 5), there was a significant relationship only between the CA-125 mean value and compliance status (p = 0.025).

Table 3: Distribution of intraoperative frozen section and final pathology results between groups

		Final pathology			Total n (%)
		Benign	Borderline	Malignant	
Intraoperative Frozen	Benign	203	4	7	214 (78.1)
	Borderline	1	13	1	15 (5.5)
	Malignant	1	2	42	45 (16.4)
Total n (%)		205 (74.8)	19 (6.9)	50 (18.3)	274 (100)

Table 4: Predictive tests for intraoperative frozen section

Adress for correspondence: Izmir Ataturk Education and Research Hospital, Karabaglar 35360 Izmir - Türkiye
e-mail: hakancokmez@hotmail.com
Available at www.actaoncologicaturcica.com

Groups	Predictive tests (%)					Prevalence (%)
	Sensitivity	Specificity	Accuracy	PPV	NPV	
Benign	99.0	84.1	95.3	94.9	96.7	74.8
Borderline	68.4	99.2	97.1	86.7	97.7	6.9
Malignant	84.0	98.7	96.0	93.3	96.5	18.3

Abbreviations: NPV, negative predictive value; PPV, positive predictive value.

Table 5: Age, mass size, and CA-125 values of agreement and no agreement for IFS and final pathology results

	Agreement	No-agreement	p-value
Age*	51.8 (13.7)	55.4 (16.4)	0.108
Mass size* (cm)	9.7 (5.7)	20 (14.4)	0.103
CA-125 value¥ (U/ml)	150.2 (34.9)	208.1 (79.4)	0.025
≤ 35¥	175 (96.7)	6 (3.3)	0.026
> 35¥	62 (88.6)	8 (11.4)	

Abbreviations: CA-125, cancer antigen 125; IFS, intraoperative frozen section.

*Mean (standard deviation).

¥Mean (standard error mean).

¥n, (%).

Table 6: Prevalence, PPV, and general accuracy of IFS in the literature

Adress for correspondence: Izmir Ataturk Education and Research Hospital, Karabaglar 35360 Izmir - Türkiye
e-mail: hakancokmez@hotmail.com
Available at www.actaoncologicaturcica.com

Study	Cases n	Prevalence %	PPV %	Accuracy %
Arora et al. (14)	292	5.1	86.7	96.2
Morton et al. (15)	277	16.3	66.7	86.6
Hashmi et al. (16)	141	4.3	83.3	98.6
Acikalin et al. (17)	282	10.3	79.3	96.5
Sukumaran et al. (18)	233	11.2	62.2	91.9

Abbreviations: IFS, intraoperative frozen section; PPV, positive predictive value.

Adress for correspondence: Izmir Ataturk Education and Research Hospital, Karabaglar 35360 Izmir - Türkiye
e-mail: hakancokmez@hotmail.com
Available at www.actaoncologicaturcica.com

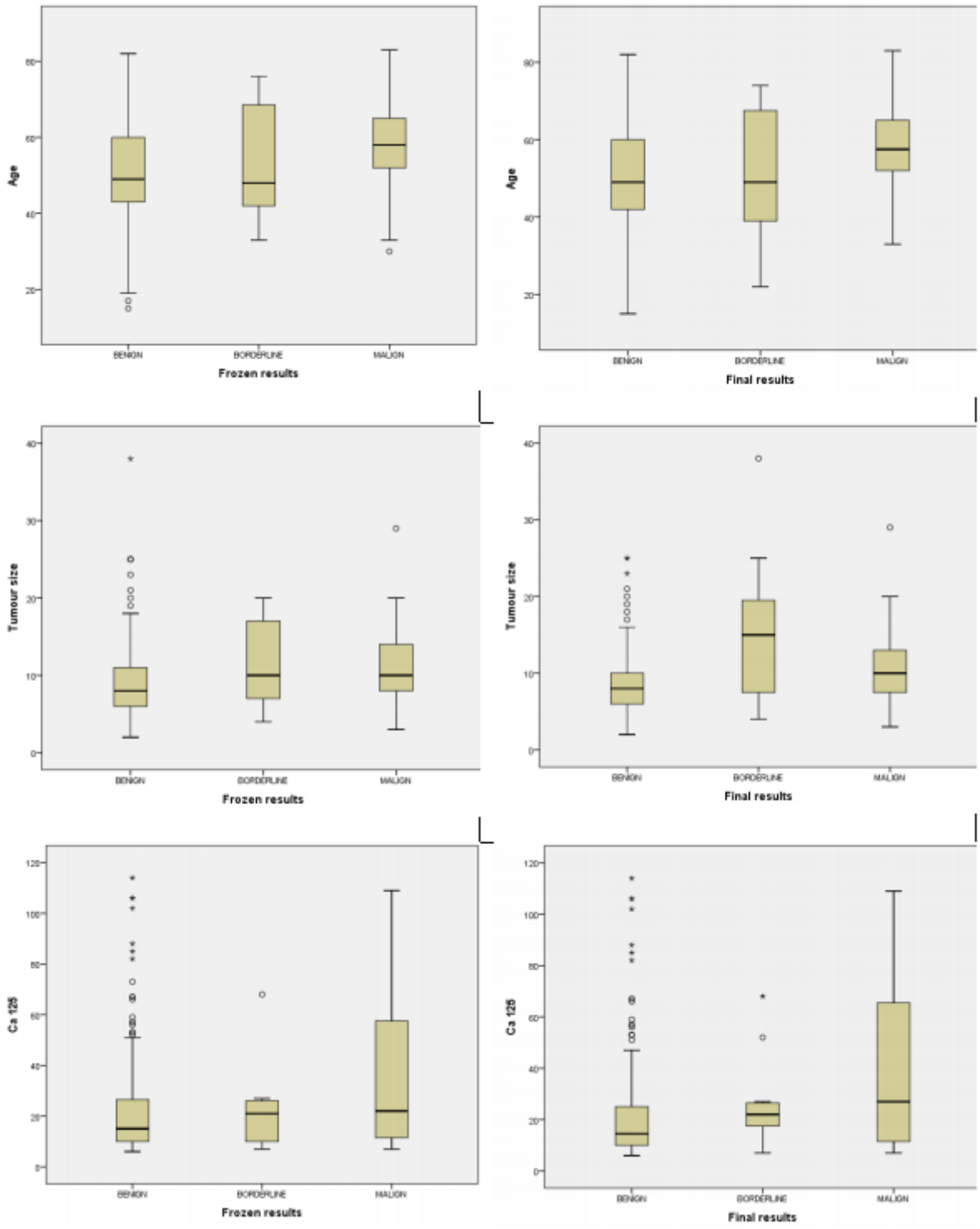

Figure 1: Distribution of age, tumor size, and cancer antigen 125 (CA-125) values among the groups.

DISCUSSION

Our study was the first to analyze the reliability of IFS and the kappa test for adnexal masses. To evaluate the conformity between the two different pathology reports (the IFS and final pathology), the kappa test was used instead of predictive statistical testing methods because it allows for the performance of calculations after extracting the results of accidental fit, which affect the overall results if not extracted, and is not affected by the prevalence ratio. Our key discovery was that the kappa value of the overall agreement showed a strong correlation according to Cohen's kappa table, despite the

Adress for correspondence: Izmir Ataturk Education and Research Hospital, Karabaglar 35360 Izmir - Türkiye
e-mail: hakancokmez@hotmail.com
Available at www.actaoncologicaturcica.com

borderline group of patients with lower prevalence, sensitivity, and PPV ratios. Huang et al. performed a study that included a meta-analysis of 13 articles and an analysis of 1577 borderline cases; they found that the sensitivity and PPV ratios (82.5% and 81.1%, respectively) were low (6). However, borderline histology can only be seen in epithelial tumors and only approximately 10% of epithelial tumors are borderline histology (10). Therefore, borderline histology has a lower prevalence rate than benign and malignant histology. This negatively affects the PPV rates of the IFS method for borderline histology.

Some studies reported in the literature have used the kappa test to evaluate the histologic grade and concordance between endometrial biopsy and final pathology results of endometrial cancer (11, 12). Similarly, the power of the endometrial biopsy to determine the histologic grade was evaluated with predictive tests; however, contradictory results were obtained because of the low prevalence of grade 2 and grade 3 histology (13). A literature search also showed that the low prevalence of borderline histology in adnexal masses causes low borderline PPV and low overall concordance rates (Table 6) (14-18). Moreover, to overcome the negative effects of borderline histology on overall concordance, a Cochrane review that compiled 38 studies involving the data of 11,181 patients with ovarian masses attempted to calculate the predictive value of IFS by including borderline cases in the benign or malignant groups (19). We think that the importance of IFS (19), which has 90% accuracy, for managing adnexal masses should be evaluated not only with predictive tests that are used to evaluate a single scan but also with the kappa test in future studies.

The missing diagnosis rates indicated a range between 2.5% and 11.9%, whereas overdiagnosis rates showed a range between 0% and 3.0% (14-18, 20). The rates of missing diagnoses and overdiagnoses in our study were also between these reported ranges. For adnexal masses evaluated with IFS, a missing diagnosis increases the risk of insufficient treatment and raises the risk of morbidity and mortality. An overdiagnosis increases the risk of not being able to perform conservative surgery, especially for premenopausal patients, possibly resulting in mortality or morbidity caused by unnecessary debulking surgery. Therefore, some studies have published histological (20), clinical (18), and laboratory (CA-125 levels) (21) data to explain disagreements between the IFS and final pathology results. Gultekin et al. (21) found that the diagnostic mismatch in cases in which CA-125 levels appeared to be ≤ 35 U/mL was seen 3.3-times more often than usual. However, we found that CA-125 levels were significantly higher in the no-agreement group. Most of the frozen analyses and all of the final pathological analyses were performed by gynecopathologists who had at least 5 years of experience, which strengthened the results of our study. However, our study was limited because it involved fewer borderline cases than other studies. Nevertheless, we think that we were able to overcome this problem by using a prevalence-independent kappa test as the statistical method. Other limitations of this study were that none of the cases had preoperative CA-125 levels, and there were differences in the expression of terms used by the pathologists who recorded comments in the medical records. In future studies related to this subject, to prevent data loss due to the aforementioned limitations, every gynecology clinic should establish a diagnostic protocol for adnexal masses, and adnexal mass cases should be prepared for surgery according to that established protocol. Additionally, to minimize the differences between pathology reports, a common template should be created with the guidance of pathologists.

In summary, the prevalence differences among benign, borderline, and malignant histology of adnexal masses affect the predictive values of IFS. When the agreement between IFS and final pathology results was evaluated with the kappa test, the strong correlation was determined to be reliable even in borderline histology cases, which had the lowest prevalence and predictive rates.

REFERENCES

1. Webb PM, Jordan SJ. Epidemiology of epithelial ovarian cancer. Best Practice Res Clin Obstet Gynaecol 2017; 41: 3-14.
2. Dora SK, Dandapat AB, Pande B, Hota JP. A prospective study to evaluate the risk malignancy index and its diagnostic implication in patients with suspected ovarian mass. J Ovarian Res 2017; 10: 55.
3. Kim K, Chung HH, Kim JW, Park NH, Song YS, Kang SB. Clinical impact of under-diagnosis by frozen section

Adress for correspondence: Izmir Ataturk Education and Research Hospital, Karabaglar 35360 Izmir - Türkiye
e-mail: hakancokmez@hotmail.com
Available at www.actaoncologicaturcica.com

examination is minimal in borderline ovarian tumors. Eur J Surg Oncol 2009; 35: 969-73.
4. Bozdag H, Guzin K, Gocmen A, Kabaca S, Usta A, Akdeniz Duran E. The diagnostic value of frozen section for borderline ovarian tumours. J Obstet Gynaecol 2016; 36: 626-30.
5. Hayran M. Saglik arastirmalari icin temel istatistik. Ankara: Omega Arastirma; 2nd edition 2018. p. 29.
6. Huang Z, Li L, Li CC, Ngaujah S, Yao S, Chu R, et al. Diagnostic accuracy of frozen section analysis of borderline ovarian tumors: a meta-analysis with emphasis on misdiagnosis factors. J Cancer 2018; 9: 2817-24.
7. Skírnisdóttir I, Garmo H, Wilander E, Holmberg L. Borderline ovarian tumors in Sweden 1960–2005: trends in incidence and age at diagnosis compared to ovarian cancer. Int J Cancer 2008; 123: 1897-901.
8. Vasconcelos I, de Sousa Mendes M. Conservative surgery in ovarian borderline tumours: A meta-analysis with emphasis on recurrence risk. Eur J Cancer 2015; 51: 620-31.
9. McHugh ML. Interrater reliability: the kappa statistic. Biochem Med 2012; 22: 276-82.
10. Gershenson DM. Management of borderline ovarian tumours. Best Pract Res Clin Obstet Gynaecol. 2017; 41: 49-59.
11. Batista TP, Cavalcanti CL, Tejo AA, Bezerra AL. Accuracy of preoperative endometrial sampling diagnosis for predicting the final pathology grading in uterine endometrioid carcinoma. Eur J Surg Oncol 2016; 42: 1367-71.
12. Martinelli F, Ditto A, Bogani K, Signorrelli M, Chiappa V, Lorusso D, et al. Accuracy of pre-operative hysteroscopic guided biopsy for predicting final pathology in uterine malignancies. J Cancer Res Clin Oncol 2017; 143: 1275-9.
13. Garcia TS, Appel M, Rivero R, Kliemann L, Wender MC. Agreement between preoperative endometrial sampling and surgical specimen findings in endometrial carcinoma. Int J Gynecol Cancer 2017; 27: 473-8.
14. Arora RS, Patel SM, Dave PS, Patel BM, Parikh CD, Poddar P. Role of Intraoperative Frozen Section in the Diagnosis of Ovarian Tumors: Experience at Gujarat Cancer and Research Institute. Indian J Gynecol Oncol. 2018;16(4), 63.
15. Morton R, Anderson L, Carter J, Pather S, Saidi SA. Intraoperative frozen section of ovarian tumors: A 6-year review of performance and potential pitfalls in an Australian tertiary referral center. Int J Gynecol Cancer 2017; 27: 17-21.
16. Hashmi AA, Naz S, Edhi MM, Faridi N, Hussain SD, Mumtaz S, Khan M. Accuracy of intraoperative frozen section for the evaluation of ovarian neoplasms: an institutional experience. World J Surg Oncol 2016; 14: 91.
17. Acikalin A, Torun G, Bagir E, Bayram F, Zeren H, Gulec U, et al. Intraoperative frozen section in ovarian neoplasms; a tertiary center experience. Turk Patoloji Derg 2014; 30: 184-8.
18. Sukumaran R, Somanathan T, Mathews A, Kattor J, Sambasivan S, Nair PR. Role of frozen section in intraoperative assessment of ovarian masses: A tertiary oncology center experience. Indian J Surg Oncol 2014; 5: 99-103.
19. Ratnavelu ND, Brown AP, Mallett S, Scholten RJ, Patel A, Founta C, et al. Intraoperative frozen section analysis for the diagnosis of early stage ovarian cancer in suspicious pelvic masses. Cochrane Database Syst Rev 2016; 3: CD010360.
20. Subbian A, Devi UK, Bafna UD. Accuracy rate of frozen section studies in ovarian cancers: A regional cancer institute experience. Indian J Cancer 2013; 50: 302-5.
21. Gultekin E, Gultekin OE, Cingillioglu B, Sayhan S, Sanci M, Yildirim Y. The value of frozen section evaluation in the management of borderline ovarian tumors. J Can Res Ther 2011; 7: 416-20.

Adress for correspondence: Izmir Ataturk Education and Research Hospital, Karabaglar 35360 Izmir - Türkiye
e-mail: hakancokmez@hotmail.com
Available at www.actaoncologicaturcica.com