Is Interpretation Differences of Radiologists an Important Issue in Breast Imaging Reporting and Data System?

Meme Görüntüleme Raporlama ve Veri Sisteminde Yorumcular Arasındaki Fark Önemli midir?

Radyologlar ve BI-RADS Kategorileme Sistemi / Radiologists and BI-RADS Categorization System

Eda Parlak[1], Cemil Gürses[1], Mustafa Yıldırım[2], Arsenal Sezgin Alikanoğlu[3], Ebru Özan[1], Ayşegül Kargı[2], Mert Köroğlu[1]
[1]Radyoloji Kliniği, [2]Tıbbi Onkoloji Kliniği, [3]Patoloji Kliniği, Antalya Eğitim ve Araştırma Hastanesi, Antalya, Türkiye

Özet

Amaç: Çalışmamızın amacı Meme Görüntüleme Raporlama ve Veri Sisteminin radyologlar arasında kullanımında uyumu araştırmak, oluşan farklılıkların nedenlerini ve bu farklılıkların en çok hangi kategorilerde gerçekleştiğinin belirlemektir. Gereç ve Yöntem: Tarama amaçlı mamografi çekilen 773 hastanın mamografileri 2 radyolog tarafından retrospektif ve birbirlerinden bağımsız olarak değerlendirildi. Tüm sonuçlar Meme Görüntüleme Raporlama ve Veri Sistemine göre değerlendirildi. Kategori 1, 2 ve 3 benign, kategori 4 ve 5 ise malignite şüpheli lezyonlar ile uyumlu olarak değerlendirildi. Gözlemcilerin benign ve malign lezyonlar arası uyumu ve tüm kategoriler arası uyumu Kappa testi ile araştırıldı. Bulgular: Çalışmaya alınan 773 hastanın 554'ünde her iki radyolog aynı kategoriyi raporlarken kalan 219 hastada farklı kategoriler raporlanmıştır. Her iki gözlemci arasındaki uyumu araştırmak için kullanılan Kappa testinin değeri istatiksel olarak anlamlı bulundu (p=0,001) ve 0.602 değeri orta derecede uyum olarak değerlendirildi. Meme Görüntüleme Raporlama ve Veri Sistemi skoruna göre benign ve malign riskte değerlendirilen hastaların mamografileri arasındaki uyum araştırıldığında benign olgular kategori 1, 2, 3 olarak kabul edilirken, kategori 4 ve 5 ise malignite şüpheli lezyonlar olarak değerlendirilmiştir. Benign ve malign olgular arasında Kappa testinin değeri istatiksel olarak anlamlı bulundu (p=0,001) ve 0.857 ile mükemmel derecede uyum olarak değerlendirildi. Her iki gözlemci arasında kategori 5 lezyonlar arasında mutlak uyum var iken en çok fark kategori 3 lezyonlarda saptandı. Tartışma: Meme Görüntüleme Raporlama ve Veri Sistemi terminolojisi memede saptanan lezyonların sınıflanması, arada oluşan karışıklıkların önlenmesi ve klinisyenlerin lezyonlara yaklaşımları açısından ortak bir dil oluşturmayı amaçlamaktadır. İstenilen düzeyde olmasa da mamografi raporlamasında belli bir standardizasyon sağlanmıştır. Ancak yorumcular arası oluşan uyumsuzluk Meme Görüntüleme Raporlama ve Veri Sisteminin yapılan çalışmalar ve görüşler doğrultusunda belirli aralıklarla güncellenmesi gerektiğini göstermektedir.

Anahtar Kelimeler

Mamografi; Meme; Radyolog

Abstract

Aim:The aim of our study is to search the reproducibility of the radiologists using Breast Imaging Reporting and Data System, to determine the reasons of differences in interpretation and find out in which categories inerterpretation differences occur mostly. Material and Method: Screening mammograms of 773 patients were evaluated retrospectively and independently by 2 radiologists. All results were evaluated according to the Breast Imaging Reporting and Data System. Category 1, 2, 3 were accepted as benign, category 4 and 5 were considered as "lesions suspicios for malignancy". Consistency of the radiologists in benign and malign lesions and also in all categories, was statistically examined by Kappa test. Results: Of the 773 patients of the study 554 were categorized identically by the 2 radiologists whereas 219 patients were categorized differently. Kappa value calculated to determine the consistency between the 2 radiologists was found statistically significant (p<0,001) and the value 0,602 was considered as moderate consistency. When consistency between mammographies of patients categorized as benign and with malignancy risk according to the Breast Imaging Reporting and Data System score is examined , benign cases were considered as category 1,2, 3 and lesions suspicious for malignancy were considered as category 4,5. Kappa value between benign and malignant cases was found statistically significant.(p<0,001) and the value of 0.857 was considered as excellent consistency. Absolute consistency was determined between the 2 radiologists in category 5 lesions whereas the most significant difference was observed in category 3 lesions. Discussion: Breast Imaging Reporting and Data System terminology aims to classify the lesions of breast and to prevent confusions in classification as well as to determine a common approach to the lesions by the clinicians.. Although not in the desired level yet, a standardization has been established in reporting mammographies. But the inconsistency between the interpreters suggest that a periodic update in the Breast Imaging Reporting and Data System according to recent studies and opinions, is neccessary.

Keywords

Mammography; Breast; Radiologist

DOI: 10.4328/JCAM.1059 Received: 02.05.2012 Accepted: 29.05.2012 Printed: 01.07.2013 J Clin Anal Med 2013;4(4): 302-304
Corresponding Author: Mustafa Yıldırım, Antalya Eğitim ve Araştırma Hastanesi, Tıbbi Onkoloji Kliniği, Varlik Mah. Kazim Karabekir Cad. Soğuksu, 07050 Antalya, Türkiye. GSM.: +905333948252 F.: +90 2422494402 E-Mail: mustafayildirim7@yahoo.com

Giriş

Meme kanseri, kadınlarda en sık rastlanan ve akciğer kanserinden sonra en çok ölümle sonuçlanan kanser tipidir [1]. Meme kanseri taramasındaki en önemli görüntüleme yöntemi mamografidir. Bu yöntemle birlikte ve tamamlayıcı olarak ultrason ve manyetik rezonans görüntüleme gibi ek yöntemler de kullanılmaktadır [2]. Tüm lezyonların ortak bir dilde anlatılabilmesi, klinisyenlerin bir sonraki aşamada neler yapabileceklerinin daha net bir dille anlayabilmeleri, mamografik bulgulardaki karışıklıkların önlenmesi için mamografide standart bir raporlama sistemine ihtiyaç duyulmuştur. Bu nedenle 1993 yılında Amerikan College of Radiology (ACR) tarafından 'Meme Görüntüleme Raporlama ve Veri Sistemi' (BI-RADS) adı verilen bir raporlama sistemi standardize edilmişti [3;4]. BI-RADS ilk yayınlandığından itibaren dört kez güncellenmiştir. Bu nedenle de BI-RADS'ın sensitivite, spesifite, ve prediktif değerleri hakkında pek çok çalışma yapılmıştır. Ancak yorumcular arasındaki uyum hakkında az sayıda çalışma bulunmaktadır [5].

Çalışmamızın amacı; aynı mamografiler için farklı radyologların BI-RADS sonuçlarını karşılaştırmak, ne kadar farklılık veya uyum olduğunu belirlemek ve özellikle bu sonuçların BI-RADS kategori 4, 5 lezyonlardaki sonuçlarını araştırmaktır. Böylece BI-RADS'ın gerçek amacına ulaşabilme oranını bir kez daha ortaya koymaktır.

Gereç ve Yöntem

Çalışmaya Antalya Kanser Erken Teşhis, Tarama ve Eğitim Merkezine Ocak 2011-Mart 2012 yılları arasında meme kanseri taraması amacıyla başvuran hastalar alındı. Mamografiler aynı ünitede çalışan ikisi de mamografi alanında deneyimli iki ayrı radyolog tarafından değerlendirildi. Değerlendirmeler değişik zamanlarda ve birbirlerinden bağımsız olarak BI-RADS kriterlerine göre yapıldı. Hastaların demografik bilgileri anket soruları aracılığı ile sorgulandı. Mamografisi iki ayrı radyolog tarafından değerlendirilmeyen hastalar çalışmaya alınmadı.

Mamografi görüntüleri GE Senographe DMR cihazı ile alındı. Tüm görüntüler rutin olarak kraniokaudal ve mediolateral oblik projeksiyonlarda çekildi. Mamografik görüntülerde single emulsiyon 18x24 cm filmler veya büyük memelerde 24x30 cm'lik filmler kullanıldı.

Her iki radyolog da filmleri BI-RADS kategorileme sistemine göre raporladı. BI-RADS kategorileme sistemine göre; Kategori 0: ek tetkik gerektirenler, 1: negatif mamografi (rutin görüntüleme), 2: benign bulgular (rutin görüntüleme), 3: büyük olasılıkla benign (kısa aralıklarla takip (6ay), 4: şüpheli anormallikler (biyopsi yapılması düşünülmeli), 5: malign olma olasılığı çok yüksek lezyonlar (gerekli işlemler yapılmalı) [6].

İstatiksel analizler SPSS 15.0 kullanılarak yapıldı. İki radyoloğun BI-RADS skorları arasındaki uyumu Cohen Kappa istatistik testi ile araştırıldı. Cohen Kappa istatistik testi iki veya daha fazla gözlem arasındaki kalitatif uyum oranını (k: Kappa) ölçmek için geliştirilmiş bir testtir. Tam uyum olması için k değerinin 1 olması gerekir. K değeri 0 olduğunda beklenilen oran şansa bağlıdır. K değeri negatif olduğunda beklenilen uyum şans ihtimalinden bile düşüktür. Mutlak bir ayrım bulunmamasına rağmen , daha önce yayınlanmış makalelerde gözlemciler arasında uyumu belirten k değeri aralıkları şöyledir; k< 0,20 ise önemsiz uyum, k: 0,21-0,40 ise minimal uyum, k: 0,41-0,60 ise orta derecede uyum, k: 0,61-0,80 ise önemli derecede uyum, k: 0,81-1,00 ise gözlemciler arası mükemmel uyum olduğunu gösterir [7].

Bulgular

Çalışmaya 773 hasta alındı. Hastaların yaş ortalaması 57.9±4.9 (51-70 yaş) olarak saptandı. Hastaların ilk menarş yaş ortalaması 13,9±1.56 (11-20 yaş), menopoz yaş ortalaması 45,6±9.47 (30-57 yaş), ilk doğum yaş ortalaması 20.7±4.7 (15-44 yaş) olarak saptandı.

Anket uygulanmasını kabul eden 211 hastanın; sigara kullanımı sorgulandığında 4'ünün (%1.9) daha önce sigara kullandığı, 13'ünün (%6.1) halen sigara kullandığı, 194'ünün (%92) sigara kullanmadığı tespit edildi. Canlı doğum sayısı medyan 3 (0-9), gebelik sayısı medyanı 4 (0-14) olarak bulundu. Oral kontraseptif kullanımı 44 (%20.8) hastada, hormon replasman tedavisi kullanımı 21 hastada (%9.9) idi. Ailesinde meme kanseri öyküsü 12 hastada (%5.7), jinekolojik kanser öyküsü 2 hastada mevcuttu.

Aynı kategoriyi 773 hastanın 554'ünde her iki radyolog raporlarken, kalan 219 hastada farklı kategoriler raporlanmıştır. Her iki radyolog tarafından şüpheli olarak kabul edilen kategori 4 ve 5 lezyonların sayısı birinci gözlemci için 12, ikinci gözlemci için 14 idi. Her iki gözlemci tarafından aynı 3 hasta kategori 5 olarak değerlendirildi. Bu hastalardan 2'si invaziv duktal karsinom, 1'i ise müsinöz karsinom tanısı aldı.

Her iki radyolog arasındaki uyumu araştırmak için kullanılan Kappa testinin değeri istatiksel olarak anlamlı bulundu. (p=0,001) ve 0.602 değeri orta derecede uyum olarak değerlendirildi. BIRADS skoruna göre benign ve malign riskte değerlendirilen hastaların mamografileri arasındaki uyum araştırılırken benign olgular kategori 1, 2, 3 olarak kabul edilirken, kategori 4 ve 5 ise maligniteşüpheli olarak değerlendirilmiştir . Benign malign olgular arası yapılan Kappa testinin değeri yine istatistiksel olarak anlamlı bulundu (p=0,001) ve 0.857 ile mükemmel derecede uyum olarak değerlendirildi. (Tablo 1)

Her iki radyoloğun BI-RADS sonuçları açısından en farklı sonuçlar kategori 3 de elde edildi. Her iki radyologda aynı sayıda kategori 5 lezyon tanımladı. Yani en çok uyum kategori 5 lezyonlarda saptandı. Her iki radyolog tarafından ortak 6 olguya kategori 4 tanısı konmuştur. Ancak kategori 4 olgularda iki radyolog arasında az sayıda da olsa fark saptanmıştır.

Tablo 1. Her iki radyologa göre BIRADS skorları

BIRADS	RADYOLOG C	RADYOLOG E
0	108 (%14)	114 (%14.7)
1	230 (%29.7)	284 (%36.7)
2	341 (%44.1)	327 (%42.3)
3	83 (%10.7)	34 (%4.4)
4	9 (%1.2)	11 (%1.4)
5	3 (%0.4)	3 (0.4)

Tartışma

Mamografi erken evre meme kanseri saptanmasında taramalarda kullanılan en etkin yöntemdir. Özellikle meme kanserinin erken saptanmasında en duyarlı bulgu olan mikrokalsifikasyonların saptanmasında doğruluğu en yüksek yöntemdir [8]. Bu nedenle mamografide kullanılan ve klinisyene lezyon hakkındaki bilgilendirmeyi yapacak olan terminoloji çok önemlidir. Termi-

nolojinin ortak olması ve klinisyenin bir sonraki aşamaya daha iyi karar verebilmesi için tüm radyologlara ve klinisyenlere yol gösterici olan BI-RADS kategorileme sistemi oluşturulmuştur. Bu zamana kadar BI-RADS ve uygulanışı konusunda pek çok çalışma yapılmıştır. Lehman ve arkadaşlarının [9] oldukça geniş bir seride ve uzun süreli yaptıkları çalışmada önceki çalışmalara oranla BI-RADS kategorileme sistemi ve lezyonlar arasında gittikçe artan bir uyum saptamışlardır. Bu da günümüzde BI-RADS kategorileme sisteminin kullanımının yaygınlaşmasını açıklamaktadır.

Birçok araştırmacı tarafından BI-RADS terminolojisinin kullanımında radyologlar arası uyum oranları test edilmiştir [6;7]. Çalışmalarda genellikle karşılaştırmalarda radyolog sayısı fazla ancak hasta sayıları az olarak incelenmiştir. Bizim çalışmamızda bu çalışmalardan farklı olarak kalabalık bir hasta grubunda iki radyolog arasındaki uyum araştırılmıştır. Benign olarak kabul edilen BI-RADS 2 ve malignite olasılığı en yüksek olan BI-RADS 5 lezyonlarda uyum en yüksek iken BIRADS 3 ve 4 lezyonlarda uyum oranı daha düşük olarak saptanmıştır. Hatta BI-RADS 5 lezyonlarda uyum en yüksek oranda bulunmuştur. Ciatto ve arkadaşlarının [4] 50 vaka ve 12 radyolog arasında yaptıkları bir çalışmada da benzer oranlar saptanmıştır Bizim çalışmamızda kanserli olguların iki radyolog arasındaki uyumu, kanserli olmayan olgulardan daha iyiydi ancak Kerlikowske ve arkadaşlarının [10] yaptığı benzer bir çalışmada bu uyum kanserli olmayan normal mamografili olgularda daha belirgin bulunmuştur. Bunun nedeni BI-RADS kategorileme sisteminde malignite yönünden şüpheli olgulara ek tetkik gerektiren olgu, yüksek olasılıkla malign veya biyopsi gerektiren olgular şeklinde değişik yaklaşımlar yapılabilmesi ve bunun da maligniteleri kategorilemede çeşitlilik oluşturduğunu belirtmişler. Ancak bizim çalışmamızda iki radyoloğun uyumu araştırıldığı için ve radyolog sayısı diğer çalışmalara göre daha az olduğundan malignitelerin uyumu daha fazla olarak saptanmıştır. Bu durumda BI-RADS kategorileme sisteminde maligniteler veya yüksek olasılıkla maligniteler konusunda radyologlar arasında ortak bir dil oluşabildiğinin göstergesidir.

Bizim çalışmamızın bir kısıtlılığınında kategori 4 lezyonları daha alt sınıflara (4a, 4b, 4c) ayırararak incelememekten kaynaklandığını düşünüyoruz. Çünkü bu sınıflamayı kullanmak benign görünüme daha yakın ancak arada kalınan kategori 4 olgularda kategori 3 yerine kategori 4a kullanılarak uyumun daha da yüksek çıkmasına neden olabilirdi. Bu konu ile ilgili Lazarus ve ark. ultrasonda kategori 4a, 4b, 4c'yi kullanarak yaptıkları çalışmada radyologlar arasında kappa değerleri oldukça yüksek bulunmuş ve kategori 4'ü subgruplara ayırmanın gereksiz biyopsileride önlediğini bildirmişlerdir [11].

BI-RADS tüm dünyada standart bir kategorileme sistemi iken kategori 4 radyologların ve klinisyenlerin kafasında soru işaretleri oluşturmaktadır. Bizim çalışmamızda ise iki radyoloğun ortak kararı kategori 4 olan 8 olgunun 1'i malign olarak saptandı. Bu sonuçta diğer çalışmalarla oldukça uyumlu bulunmuştur. Kategori 4 olgular arası uyumsuzluğun nedeninin ise radyoloğun bilgi ve deneyiminden çok diğer etkenlere bağlıdır. Bunların içinde en önemlisi meme dansitesidir. Mamografinin duyarlılığı skleroze veya liposkleroze memelerde azalmaktadır. Bu durum lezyon tanımlamalarında ve BIRADS kategorilemelerinde farklılıklara neden olmaktadır. Radyologlar arasında meme dansitesine yönelik BİRADS ile ilgili yapılan bazı çalışmalarda uyumun oldukça yüksek olduğu saptanmıştır [12;13].

Daha önce yapılan ve lezyonların morfolojik özellikleri, mikrokalsifikasyon görünümleri veya kitlelerin değişik görünme şekilleri ile ilgili yapılan çalışmalarda radyologlar arası uyumun BI-RADS ile yapılan çalışmalarda saptanan uyumdan çok daha az olduğu bildirilmiştir [4;7]. Bu da BI-RADS'ın radyologlar arasında ortak bir dil oluşturma konusunda ne kadar başarılı olduğunu göstermektedir.

Sonuç olarak BI-RADS kategorileme sistemi mamografi raporlarını standardize etmek, mamografi yorumlarında karışıklığı önlemek, sonuçların yol göstericiliğini kolaylaştırmak ve ortak bir terminoloji oluşturmak amacıyla geliştirilmiştir ve bu konuda da oldukça uzun bir yol katetmiş durumdadır. Ancak yorumcular arası oluşan uyumsuzluk bize bu konuda yapılan çalışmaların gerekliliğini ve BI-RADS kategorileme sisteminin bu görüşler doğrultusunda belirli aralıklarla güncellenmesi gerektiğini göstermektedir.

Kaynaklar

1. Siegel R, Naishadham D, Jemal A. CA Cancer J Clin. 2012;62(1):10-29.
2. Parlak E, Yıldırım M, Ozkur A, Andıç Ç, Bayram M. A comparation of imaging methods in the evaluation breast patologies. Gaziantep Tıp Dergisi. 2011;17:114-20.
3. Eberl MM, Fox CH, Edge SB, Carter CA, Mahoney MC. BI-RADS classification for management of abnormal mammograms. J Am Board Fam Med. 2006;19:161-4.
4. Ciatto S, Houssami N, Apruzzese A, Bassetti E, Brancato B, Carozzi F, et al. Reader variability in reporting breast imaging according to BI-RADS assessment categories. Breast. 2006;15:44-51.
5. Antonio AL, Crespi CM. Predictors of interobserver agreement in breast imaging using the Breast Imaging Reporting and Data System. Breast Cancer Res Treat. 2010;120:539-46.
6. Berg WA, Campassi C, Langenberg P, Sexton MJ. Breast Imaging Reporting and Data System: inter- and intraobserver variability in feature analysis and final assessment. AJR Am J Roentgenol. 2000;174:1769-77.
7. Gülsün M, Demirkazık FB, Köksal A, Arıyürek M. Meme mikrokalsifikasyonlarının BI-RADS kriterlerine göre değerlendirilmesi ve yorumcular arasındaki uyumun araştırılması. Meme Radyolojisi. 2002;8:358-63.
8. Baker JA, Kornguth PJ, Floyd CE Jr. Breast imaging reporting and data system standardized mammography lexicon: observer variability in lesion description. AJR Am J Roentgenol. 1996;166:773-8.
9. Lehman C, Holt S, Peacock S, White E, Urban N. Use of the American College of Radiology BI-RADS guidelines by community radiologists: concordance of assessments and recommendations assigned to screening mammograms. AJR Am J Roentgenol. 2002;179:15-20.
10. Kerlikowske K, Grady D, Barclay J, Frankel D. S, Ominsky H. S, Sickles A.E. Variability and Accuracy in Mamographic Interpretation Using the American College of Radiology Breast Imaging and Data System Journal of the National Cancer Institute. 1998;23:1801-9.
11. Lazarus E, Mainiero MB, Schepps B, Koelliker SL, Livingston LS. BI-RADS lexicon for US and mammography: interobserver variability and positive predictive value. Radiology. 2006;239:385-91.
12. Ciatto S, Houssami N, Apruzzese A, Bassetti E, Brancato B, Carozzi F, et al.Categorizing breast mammographic density: intra- and interobserver reproducibility of BI-RADS density categories. Breast. 2005;14:269-75.
13. Ooms EA, Zonderland HM, Eijkemans MJ, Kriege M, Mahdavian Delavary B, Burger CW, et al. Mammography: interobserver variability in breast density assessment. Breast. 2007;16:568-76.