

유방 밀도 평가: 영상의학과 의사의 육안 평가와 Volpara 유방 밀도의 비교

고수연 · 김은경 · 김민정 · 문희정

연세대학교 의과대학 영상의학과

목적: 영상의학과 의사가 판정한 유방조직 밀도와 Volpara 소프트웨어의 유방조직 밀도의 일치도를 비교하고 Volpara 소프트웨어의 밀도 측정이 오류를 보이는 원인에 대해 알아 보고자 한다.

대상 및 방법: 2011년 12월 20일부터 2011년 12월 24일까지 본원에서 유방촬영술을 시행한 116명의 환자를 대상으로 하였다. 유방 촬영술의 유방밀도를 7명의 영상의학과 의사가 미국방사선의학회의 유방영상 판독 및 데이터체계 (Breast Imaging Reporting and Data System: BI-RADS, American College of Radiology)에 의거하여 판독하였고, Volpara 소프트웨어에서 측정된 유방 밀도 결과와 비교 분석하였다. 두 진단 방법의 일치도를 평가하기 위하여 가중된 카파 분석 (Weighted Kappa test)을 이용하였다. Volpara 유방 밀도 측정의 오류가 생겼을 때 그 원인에 대해 분석하였다.

결과: 영상의학과 의사의 육안적 평가와 Volpara 유방 밀도 결과의 일치도는 중증도 일치도를 보였다 (카파 값 =0.428). 116명의 환자 중 5명 환자 (4.3%)에서 Volpara 유방 밀도 측정 오류가 발생하였다. 오류가 발생한 군과 그렇지 않은 군 사이에 총 유방 볼륨과 유방 밀도의 유의한 차이가 있었다 (p=0.0301, p=0.0067). ROC 곡선 분석에서 총 유방 볼륨 367.9를 기준으로 하였을 때 민감도 100%, 특이도 63%였다.

결론: 영상의학과 의사와 Volpara의 유방 밀도는 중증도의 일치도를 보였다. Volpara 유방 밀도 측정은 4.3%에서 오류가 나타났고 오류가 발생한 군과 그렇지 않은 군 사이에 총 유방 볼륨과 유방 밀도에 유의한 차이가 있었다.

Index words: Mammographic breast density; Volpara; Breast Imaging Reporting and Data System

서 론

유방 밀도는 유방촬영 영상에서 유방 전체 면적 중 섬유유선 조직이 차지하는 비율로 정의할 수 있다. 미국방사선의학회 (American College of Radiology: ACR)의 유방영상 판독 및 데이터체계 (Breast Imaging

Reporting and Data System: BI-RADS, 이하 BI-RADS)에서 제시한 기준에 따라 정성적인 유방 밀도를 평가하여 4단계 범주로 보고한다(1). 범주 3, 4는 고밀도 유방으로 분류되며 고밀도 유방은 유방암의 위험인자일 뿐 아니라 유방암의 위음성의 요인이 된다(2, 3). 범주 4의 치밀유방인 경우 유방촬영술의 민감도는 30-48%로 감소한다(4, 5). 따라서, 정확한 유방 밀도의 판정은 환자의 유방암 위험도 평가뿐 아니라 환자의 위음성을 예측에 중요하다.

유방 촬영 영상에서 유방 밀도를 측정하는 방법을 크게 정성적 방법과 정량적 방법인 두 가지로 분류할 수 있다(6). 정성적 방법은 영상의학과 의사가 판독하는 것으

통신저자: 문희정
(120-752) 서울특별시 서대문구 신촌동 134
연세의료원건물 영상의학과
Tel. (02)2228-7400, Fax. (02)393-3035
E-mail: artemis4u@yuhs.ac

로 전체 유방에서 섬유유선 조직이 차지하는 비율을 육안으로 보고 범주에 따라 분류하는 것으로 Wolfe 분류법, Tabar 분류법, ACR BI-RADS 등이 있다(7). 정성적인 방법 중 하나인 ACR BI-RADS의 4 범주가 현재 널리 쓰이고 있으나 주관적이며, 판독자간, 판독자내 변이가 많고 재현성이 낮은 단점이 있다(1, 8). 정량적인 방법은 정성적인 방법과 달리 유방 볼륨 중 유방 밀도가 차지하는 비율을 계산하므로 더 객관적이고 재현성을 높일 수 있다(9, 10). 유방 밀도의 정확한 측정과 재현성은 유방암 위험 예측 모델 또는 역학 연구에서 매우 중요하다(8). 객관적이고 재현성이 높은 표준화된 유방밀도 측정 방법이 있어야 유방 밀도와 유방암 발병 위험 모델이나 유방 밀도에 영향을 줄 수 있는 변수(약물)의 영향에 대한 임상 실험 연구의 유용성이 높아진다(8). 또한 유방암 위험도 측정과 함께 고위험군 여성에서 검사 주기 및 검진 방법 결정에 사용될 수 있으므로, 맞춤형 진단과 치료를 표방하는 향후 의료의 발전 방안에도 뜻을 같이 할 수 있겠다(11). Volpara 소프트웨어는 볼륨 기반의 유방 밀도 측정법 중 하나로 양측 총 유방 볼륨(Total breast tissue volume, cm³)과 섬유유선 조직의 볼륨(fibroglandular tissue volume, cm³)의 상대적 비율(%)을 자동적으로 제공한다.

본 연구는 영상의학과 의사가 분류한 ACR BI-RADS에 의거한 유방밀도의 범주와 Volpara 소프트웨어에서 측정된 결과가 얼마나 일치하는지를 알아보고자 한다. 또한 Volpara 유방 밀도 측정에 오류가 나는 경우 어떠한 변수가 영향을 줄 수 있는지 알아보고자 하였다.

대상 및 방법

1. 대상 환자

2011년 12월 20일부터 2011년 12월 24일까지 본원에서 유방 촬영술을 시행한 환자는 총 122명이었다. 그 중 Volpara 유방 밀도의 결과가 없는 환자 6명을 제외하고 최종적으로 116명 환자가 본 연구에 포함되었다.

2. 영상검사

유방 촬영은 2대의 디지털 유방촬영기를 이용하였다(Lorad/Hologic Selenia full-field digital mammography system (Lorad/Hologic, Danbury, CT, USA) and the General Electric senograph digital mammography system (General Electric Medical Systems, Milwaukee, WI, USA)). 유방 촬영술 판독은 7명의 영상의학과 전문의가 ACR BI-RADS의 권고에 따라 유방 구성(breast composition)을 평가하였

다. 유방 실질 유형을 지방형(almost entirely fatty, <25% glandular), 유선산재형(scattered fibroglandular, 25-50% glandular), 중등도치밀형(heterogeneously dense, 51-75% glandular), 고등도치밀형(extremely dense, >75% glandular)의 4 범주로 나누어 평가하였다.

Volpara (Version 1.5.1, Matakina Technology, LTD, Wellington, New Zealand)를 이용한 유방 밀도 측정은 섬유유선 조직 볼륨을 총 유방 볼륨으로 나눈 유방 밀도(volumetric breast density, %)가 측정되었다. 좌우 유방 각각에서 계산된 유방 밀도의 평균이 한 환자의 Volpara 유방 밀도로 측정되었고, 유방밀도 단계는 1 (0-4.5%), 2 (4.5-7.5%), 3(7.5-15.5%), 4 (>15.5%)의 단계로 분류된다. 각각의 단계는 ACR BIRADS의 지방형, 유선산재형, 중등도치밀형, 고등도치밀형에 해당한다. 육안으로 평가한 유방 밀도와 Volpara 유방 밀도 측정값의 일치도를 평가하기 위해 가중된 카과 분석을 이용하였다. 유방 밀도 범주 1, 2를 저밀도로 3, 4를 고밀도 유방으로 분류하여 영상의학과 의사와 Volpara 유방 밀도간의 일치도를 카과 분석을 통해 비교하였다. 카과 값(k value)이 0.2 이하는 불량(slight agreement), 0.21 - 0.40는 보통(fair agreement), 0.41 - 0.60는 중등도(moderate agreement), 0.61 - 0.80은 우수(substantial agreement), 0.81 - 1.00는 아주 우수(almost perfect agreement)로 분류하였다.

3. 데이터 및 통계 분석

통계적 분석은 연속 변수들에서 Mann-Whitney test, 명목 변수들에서는 Fisher's exact test를 이용하였다. Volpara 유방 밀도의 오류가 발생한 집단과 그렇지 않은 집단의 총 유방 볼륨, 섬유유선 조직 볼륨, 유방 밀도를 비교하기 위해 Mann-Whitney test를 이용하였다. 오류가 발생한 유방의 볼륨은 0으로 측정되므로 반대쪽 유방 볼륨만을 이용하여 계산하였고 오류가 없는 유방의 볼륨은 양쪽 유방 볼륨의 평균으로 비교하였다. 오류 발생에 기여하는 인자에 대해 분석하기 위해 환자군을 수술을 한 군과 수술을 하지 않은 군으로 나누었다. 유방전절제술을 시행한 경우 수술을 시행하지 않은 유방의 볼륨만을 이용하여 분석을 하였고 수술을 하지 않은 군에 포함시켰으며, 부분절제술을 시행한 환자만을 수술을 시행한 군에 포함시켰다. 두 집단 사이에 Volpara 유방 밀도 측정 오류 발생률의 의미 있는 차이가 있는지 분석하기 위해 Fisher's exact test를 이용하였다. P 값이 0.05 미만인 경우 통계학적으로 유의하다고 판단하였

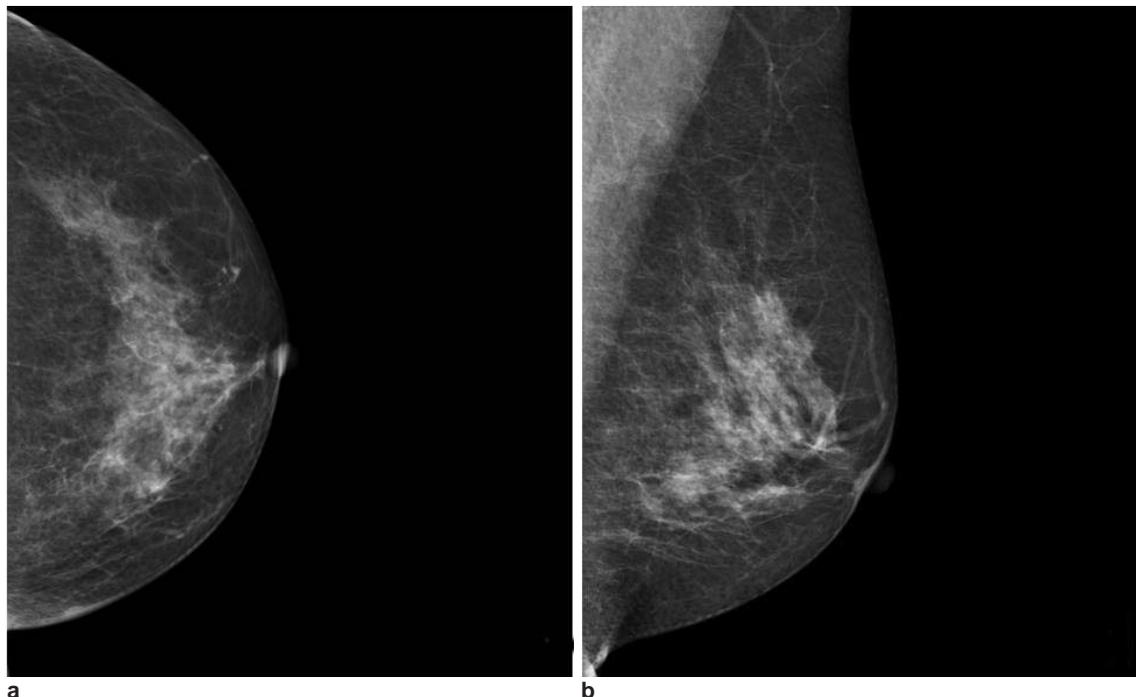
다. Volpara 유방 밀도 측정에 오류가 발생하게 되는 총 유방 볼륨을 알아보는 적절한 역치를 찾기 위해 receiver operating characteristics (ROC) 곡선 분석을 이용하였다. 분석은 MedCalc (version 11.5.0.0, MedCalc Software, Mariakerke, Belgium)을 이용하였다.

결 과

116명의 환자 (25-74세, 평균 51세) 중 84명은 수술을 받지 않았고, 24명은 유방 전절제, 8명은 한 쪽 유방의 부분 절제를 받았다. 116명의 평균 유방 밀도 (%), 섬유유선 조직 볼륨/총 유방 볼륨)는 17.04%, 평균 총 유방 볼륨은 513.77 mL, 평균 섬유유선 조직 볼륨은 52.7

mL 였다. 영상의학과 의사가 판독한 유방밀도와 Volpara 유방 밀도 측정값의 일치도는 카파 값이 0.428로 중증도 일치의 결과를 보였다 (Fig. 1). 저밀도 (유방 밀도 범주 1, 2)와 고밀도 (유방 밀도 범주 3, 4)로 분류 시 영상의학과 의사와 Volpara 유방 밀도 측정값의 일치도는 카파 값이 0.554로 중증도 일치의 결과를 얻었다.

Volpara 소프트웨어의 유방 밀도 측정의 오류는 116명 환자 중 5명인 4.3%에서 나타났으며, 총 유방 볼륨과 섬유유선 조직 볼륨이 0으로 표시되었다. 8명의 유방 부분절제술을 받은 환자 중 1명 (12.5%)이 Volpara 유방 밀도 측정 시 총 유방 볼륨과 섬유유선 조직 볼륨이 0으로 표시되는 오류가 발생했으며 수술을 받지 않거나 전 절제술을 받은 108명 중 4명 (3.7%)에서 오류가 발생하였고 통계학적 차이는 없었다 ($p=0.3050$).



Volpara v1.4
Breast Density Assessment

Patient Name
Patient ID
Patient DOB 02/16/1956
Accession # 1106127078
Study Date 12/14/2011

Factor	Right	Left
Fibroglandular Tissue Volume (cm3)	0.0	50.3
Breast Tissue Volume (cm3)	0.0	633.2
Breast Density (%)	0.0	7.9

Volpara Density Grade (tm) 2

Fig. 1. A 55-year-old woman performed mammography after right total mastectomy and left mammogram (a, b) was done. Breast density was classified as grade 3 by a radiologist but Volpara imaging software (c) interpreted grade 2 parenchymal density.

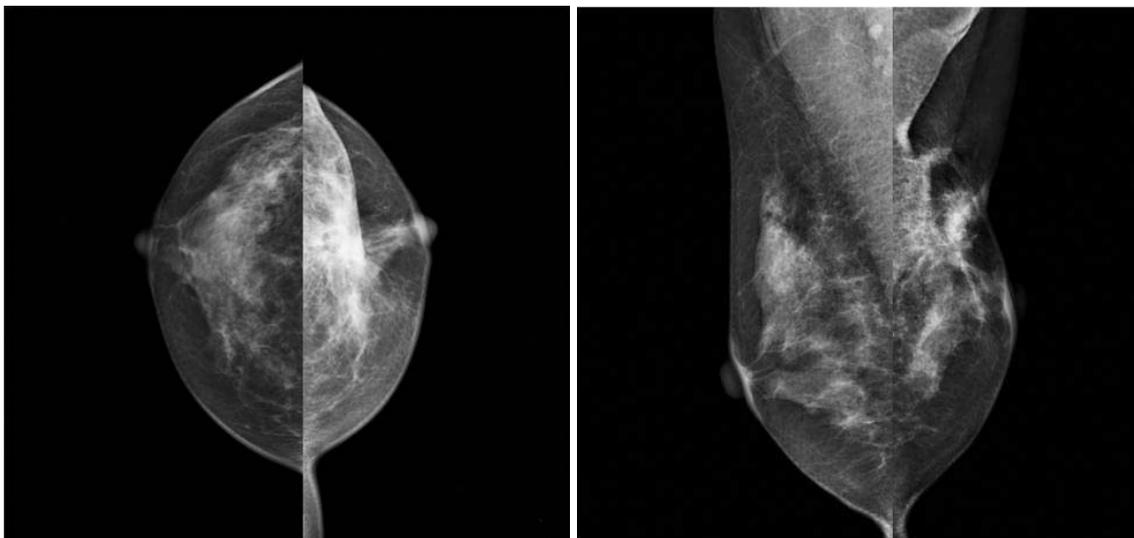
Volpara 소프트웨어에서 측정된 총 유방 볼륨은 오류가 발생한 군(median 330.7 mL)에서 그렇지 않은 군 (median 451 mL)에서보다 더 작았다 ($p=0.0301$ Mann-Whitney U test) (Table 1). 오류 발생군의 유방 섬유유선 조직 볼륨의 중간값은 66.7 mL, 발생하지 않은 유방은 41.3 mL로 통계학적 유의성은 없으나 오류가 발생하지 않은 유방의 섬유유선 조직 볼륨이 상대적

으로 적었다 ($p=0.5819$). 오류가 발생한 유방의 평균 유방밀도의 중간값은 18 mL, 발생하지 않은 유방의 중간값은 10.9 mL로 유의하게 오류가 발생한 유방에서 높았다 ($p=0.0067$) (Fig. 2). ROC 분석에서 Volpara 유방 밀도 범주의 오류를 발견하는데 있어 유용한 총 유방 볼륨의 컷오프는 367.9 mL 였다. (area under the ROC curve; 0.787, 95% confidence interval 0.702-

Table 1. Comparison on Total Breast Volume, Fibroglandular Volume and Breast Density Using Mann-Whitney Test between Groups With and Without Error

	Error (+) (n = 5)	Error (-) (n = 111)	p-value
Median Total Breast Tissue Volume (mL)	330.7	451	0.0301
(Range)	(86.2-367.9)	(88.5-6553.6)	
Median Fibroglandular Tissue Volume (mL)	66.7	41.3	0.5819
(Range)	(14.9-77.1)	(10.9-140.9)	
Breast Density (%)	18	10.9	0.0067
(Range)	(14.5-22.7)	(3.5-148.0)	

p-value < 0.05 – statistically significant



a **b**

```

Volpara v1.4
Breast Density Assessment

Patient Name
Patient ID
Patient DOB 10/17/1951
Accession # 1107009745
Study Date 12/19/2011

Factor          Right    Left
Fibroglandular Tissue Volume (cm3)  36.7    0.0
Breast Tissue Volume (cm3)          251.7   0.0
Breast Density (%)                   14.5    0.0

Volpara Density Grade (tm)          3
    
```

c

Fig. 2. A 60-year-old woman with left partial mastectomy underwent mammogram. Breast density was interpreted as grade 3 by a radiologist and Volpara (a, b). Volpara (c) showed only right sided volume and density. The volume and density of left breast were 0, representing technical error. Small breast tissue volume due to postoperative change might be the cause of error.

0.858, 민감도: 100%, 특이도: 63.06%).

고 찰

유방 밀도는 유방암의 위험인자이다(12). 한 메타 분석 (meta-analysis)에서 높은 유방촬영 유방밀도 (>75%)를 가진 여성의 경우 낮은 유방밀도(<10%)를 가진 여성에 비해 유방암의 위험도가 증가한다고 보고하였다 (odds ratio, 4.7; 95% confidence interval 3.0 - 7.4) (11). 또한 높은 유방밀도는 유방촬영술의 정확한 판독을 어렵게 하고, 유방암을 놓칠 확률과 위양성율이 모두 높다(13). 유방 밀도는 유방암 위험 평가 연구 및 고밀도 유방을 가진 환자에서의 스크리닝 방법 제시에 있어 분명히 중요한 요소이다(14). ACR BI-RADS는 현재 임상에서 많이 쓰이는 보고체계이나 판독자간, 판독자내 상당한 변이가 있고 재현성이 낮다(9, 15). 객관적이고 재현성이 높을 뿐만 아니라, 유방암 위험인자 관련 연구 등에 널리 사용할 수 있는 정량적 유방 밀도 분석 방법이 필요하였고 최근 유방 밀도를 정량화하는 소프트웨어들이 개발 되었다(9, 11, 16). 이 중 하나인 Volpara 소프트웨어를 이용하여 유방밀도에 대한 연구를 하였다.

Volpara 소프트웨어를 이용하여 유방밀도 분석 시 영상의학과 의사의 정성적 분류와 Volpara의 정량적 분석은 카파값이 0.428로 중증도의 일치도를 보였다. 범주 1, 2를 저밀도로 3, 4를 고밀도로 분류하여 시행한 일치도 분석 시 카파값이 0.554로 중등도 일치도를 보였다. Tagliafico 등의 연구는 우리의 결과와는 달리 아주 우수한 일치도를 보였다(7). 우리의 연구는 7명의 영상의학과 의사가 분류한 범주를 이용하였고, 7명의 판독자간 변이를 분석하지 않은 것이 그 이유로 생각된다. 향후 많은 환자를 포함한 전향적 연구가 필요할 것으로 생각된다.

본 연구에서 Volpara 소프트웨어의 유방 밀도 측정의 오류는 116명 환자 중 5명인 4.3%에서 나타났으며, 총 유방 볼륨과 섬유유선 조직 볼륨이 0으로 표시되었다. Volpara 소프트웨어에서 측정된 총 유방 볼륨은 오류가 발생한 군 (중간값 330.7 mL)에서 그렇지 않은 군 (중간값 451 mL)에서보다 더 작았다 ($p=0.0301$ Mann-Whitney test). 오류 발생군의 유방 섬유유선 조직 볼륨의 중간값은 66.7 mL, 발생하지 않은 유방은 41.3 mL로 통계학적 유의성은 없으나 오류가 발생하지 않은 유방의 섬유유선 조직 볼륨이 상대적으로 적었다 ($p=0.5819$). 오류가 발생한 유방의 평균 유방밀도의 중간값은 18 mL, 발생하지 않은 유방의 중간값은 10.9

mL로 유의하게 오류가 발생한 유방에서 높았다 ($p=0.0067$). 유방 밀도가 높은 경우 지방조직만을 포함하는 신뢰할 만한 교정(calibration) 위치를 얻는 것이 매우 어렵다(19). 교정의 전제 조건은 압박된 유방 두께의 정확한 평가이다. 압박된 상태에서 균등한 유방 두께를 얻기 위해 곡선부분을 처리하는 과정에서 유방 볼륨이 작을 경우 오류에 더 민감할 것이다. 유방을 압박 시 유방 두께를 유방 압박 패들이 단단하나 유연하나에 따라 유방 변연부에서 흉벽까지 유방 두께가 최대 2 cm 까지 달라질 수 있다. 이러한 유방 두께의 변이는 볼륨 기반 유방밀도 측정법의 오류의 원인 중 하나이며 압박 조건에 따라 균등한 유방 두께 측정에 한계가 있을 것이다(20). 이러한 이유가 유방 밀도가 높은 경우에서 오류가 발생하는데 기여하였을 가능성이 있다. 소프트웨어를 이용한 유방밀도 측정법은 유방 테두리를 감지하는 과정에서 오류가 생길 수 있으며 유방 자동 분할화 작업은 수동 작업에 비해 오류에 더 민감할 수 있다. 특히 유방 밀도가 높을 경우 자동 분할화 수행 능력이 더 떨어질 것으로 생각한다. 총 유방 볼륨이 작을수록 상대적으로 섬유유선 조직이 차지하는 비율이 증가하므로, 본 연구에서 총 유방 볼륨이 작은 군에서 오류 발생이 유의하게 높았던 결과는 높은 유방 밀도와 연관지어 설명할 수 있다.

본 연구는 몇가지 제한점을 가지고 있다. 첫째, 소수의 환자를 대상으로 한 후향적 연구이다. 연구의 결과를 일반화 하기 위해서는 더 큰 집단을 대상으로 한 전향적 연구가 필요하다. 둘째, 7명의 영상의학과 의사가 판독한 유방밀도와 Volpara 소프트웨어를 비교하였으나, 7명 판독자간 변이는 측정하지 못하였다. 이것이 다른 연구보다 낮은 일치도를 보이는 결과의 원인일 수 있다.

결론적으로, 향후 많은 대상군을 포함한 전향적 연구가 필요하나 Volpara 소프트웨어의 정량적 유방밀도 분석과 영상의학과 의사의 정성적 분석은 중증도의 일치도를 보였다. Volpara 소프트웨어는 전체 유방 볼륨이 작거나 유방 밀도가 높은 경우에는 기술적인 오류가 발생할 수 있다.

참 고 문 헌

1. Balleyguier C, Ayadi S, Van Nguyen K, Vanel D, Dromain C, Sigal R. BI-RADS classification in mammography. *Eur J Radiol* 2007;61:192-194
2. Ducote JL, Molloy S. Quantification of breast density with dual energy mammography: an experimental feasibility study. *Med Phys* 2010;37:793-801
3. Boyd NF, Guo H, Martin LJ, et al. Mammographic density and the risk and detection of breast cancer. *N Engl J Med* 2007;356:227-236
4. Kolb TM, Lichy J, Newhouse JH. Comparison of the performance

- of screening mammography, physical examination, and breast US and evaluation of factors that influence them: an analysis of 27,825 patient evaluations. *Radiology* 2002;225:165-175
5. Mandelson MT, Oestreicher N, Porter PL, et al. Breast density as a predictor of mammographic detection: comparison of interval- and screen-detected cancers. *J Natl Cancer Inst* 2000;92:1081-1087
 6. Yaffe MJ. Mammographic density. Measurement of mammographic density. *Breast Cancer Res* 2008;10:209
 7. Tagliafico A, Tagliafico G, Tosto S, et al. Mammographic density estimation: comparison among BI-RADS categories, a semi-automated software and a fully automated one. *Breast* 2009;18:35-40
 8. Garrido-Esteba M, Ruiz-Perales F, Miranda J, et al. Evaluation of mammographic density patterns: reproducibility and concordance among scales. *BMC Cancer* 2010;10:485
 9. Gao J, Warren R, Warren-Forward H, Forbes JF. Reproducibility of visual assessment on mammographic density. *Breast Cancer Res Treat* 2008;108:121-127
 10. Buchbinder SS, Leichter IS, Lederman RB, et al. Computer-aided classification of BI-RADS category 3 breast lesions. *Radiology* 2004;230:820-823
 11. Martin KE, Helvie MA, Zhou C, et al. Mammographic density measured with quantitative computer-aided method: comparison with radiologists' estimates and BI-RADS categories. *Radiology* 2006;240:656-665
 12. Ding J, Warren R, Warsi I, et al. Evaluating the effectiveness of using standard mammogram form to predict breast cancer risk: case-control study. *Cancer Epidemiol Biomarkers Prev* 2008;17:1074-1081
 13. Sala E, Warren R, McCann J, Duffy S, Day N, Luben R. Mammographic parenchymal patterns and mode of detection: implications for the breast screening programme. *J Med Screen* 1998;5:207-212
 14. van Gils CH, Otten JD, Hendriks JH, Holland R, Straatman H, Verbeek AL. High mammographic breast density and its implications for the early detection of breast cancer. *J Med Screen* 1999;6:200-204
 15. Glide-Hurst CK, Duric N, Littrup P. A new method for quantitative analysis of mammographic density. *Med Phys* 2007;34:4491-4498
 16. Yaffe MJ, Boyd NF, Byng JW, et al. Breast cancer risk and measured mammographic density. *Eur J Cancer Prev* 1998;7 Suppl 1:S47-55
 17. Lokate M, Kallenberg MG, Karssemeijer N, Van den Bosch MA, Peeters PH, Van Gils CH. Volumetric breast density from full-field digital mammograms and its association with breast cancer risk factors: a comparison with a threshold method. *Cancer Epidemiol Biomarkers Prev* 2010;19:3096-3105
 18. Byng JW, Yaffe MJ, Jong RA, et al. Analysis of mammographic density and breast cancer risk from digitized mammograms. *Radiographics* 1998;18:1587-1598
 19. van Engeland S, Snoeren PR, Huisman H, Boetes C, Karssemeijer N. Volumetric breast density estimation from full-field digital mammograms. *IEEE Trans Med Imaging* 2006;25:273-282
 20. Mawdsley GE, Tyson AH, Peressotti CL, Jong RA, Yaffe MJ. Accurate estimation of compressed breast thickness in mammography. *Med Phys* 2009;36:577-586
 21. McCormack VA, Highnam R, Perry N, dos Santos Silva I. Comparison of a new and existing method of mammographic density measurement: intramethod reliability and associations with known risk factors. *Cancer Epidemiol Biomarkers Prev* 2007;16:1148-1154
 22. Alonzo-Proulx O, Packard N, Boone JM, et al. Validation of a method for measuring the volumetric breast density from digital mammograms. *Phys Med Biol* 2010;55:3027-3044

Mammographic Density Estimation: Comparison Between Radiologist's Visual Assessment and Volpara Breast Density

Su Yeon Ko, MD, Eun-Kyung Kim, MD, Min Jung Kim, MD, Hee Jung Moon, MD

Department of Radiology, Yonsei University College of Medicine

Purpose: To retrospectively compare Volpara mammographic density estimation with the density by radiologists based on the Breast Imaging Reporting and Data System (BI-RADS) density classification.

Materials and Methods: Materials and methods: From twentieth December 2011 to twenty fourth December 2011, 116 patients aged 25-74 years who had mammography with Volpara density grade were included in this study. We reviewed the patient's breast density grade of BI-RADS categories by radiologists and Volpara. Weighted kappa statistic was used in measuring agreement between radiologists and Volpara imaging software. Mann Whitney test was used for comparison total breast volume (mL), fibroglandular tissue volume (mL), and degree of breast density (%) in group with technical error of Volpara imaging software and group without error.

Results: Overall agreement of breast density between the radiologist's BI-RADS and Volpara imaging software was moderate ($k=0.428$). Five (4.3%) patients had technical error of Volpara imaging software in breast density measurements. Statistical analysis showed a significant difference in total breast tissue volume (mL) and degree of breast density (%) when comparing group with technical error of Volpara software and group without technical error of that method ($p=0.0301$, $p=0.0067$).

Conclusion: The agreement of radiologists and Volpara imaging software was moderate (kappa value = 0.428). Small total breast volume (mL) and high degree of breast density (%) were causes of the Volpara software error.

Index words: Mammographic breast density; Volpara; Breast Imaging Reporting and Data System

Corresponding author: Hee Jung Moon, M.D.