

치밀유방 여성에서 MRI 검진: DENSE 및 EA1141 영상시험의 결과와 해석

문 우 경

서울대학교 의과대학 영상의학교실

서 론

유방암 검진의 표준 방법인 유방촬영은 지방 조직보다 섬유유선 조직의 비율이 높은 치밀유방 여성에서 민감도가 떨어지고 간격암 발생률이 높다 (1). 또한 치밀유방은 유방암의 위험인자로 치밀형 유방 조직의 여성은 지방형 유방 조직의 여성에 비해 유방암 위험도가 2-5배 높다 (1, 2). 미국의 유방밀도 고지법률은 이러한 유방촬영 검진의 문제점을 반영한 것이며 이를 계기로 치밀유방 여성에서 추가검진 영상검사 방법과 선택에 필요한 정보 제공의 필요성이 증가하고 있다 (3, 4). 현재 치밀유방 여성에서 유방촬영에 추가 검진 목적으로 사용 가능한 영상검사로는 초음파, 디지털 유방 단층촬영 (digital breast tomosynthesis, DBT)과 MRI 등이 있다. 유방초음파가 접근도가 높아 국내외에서 가장 많이 사용되고 있지만 암 발견율이 낮고 양성 조직검사와 단기 추적검사 등 위양성률이 높은 것이 제한점이다 (5). 3차원 유방촬영인 DBT는 디지털 유방촬영보다 암발견율은 높고 소환율은 낮아서 추가 검사로서의 역할 보다는 디지털 유방촬영을 대체하는 표준 유방촬영방법으로 정착되고 있다. 하지만 극도 치밀유방 여성에서 DBT의 민감도는 유방초음파보다 유의하게 낮다 (6). 유방 MRI는 유방 밀도와 상관없이 모든 유방암 위험도의 여성에서 암 발견에 민감도가 가장 높지만 위양성률도 비교적 높고 검사 비용이 고가인 것이 검진

에 광범위하게 사용하는데 문제이다 (7, 8). 이러한 제한점 때문에 미국에서도 고위험 여성의 5% 미만에서만 실제 MRI 검진을 받은 것으로 나타났다 (9). 2014년 독일의 Kuhl이 제안한 단축 (abbreviated breast, AB) MRI는 검사 시간과 판독 시간을 줄이면서도 진단 성적은 전체 프로토콜 MRI와 비슷하다고 알려지면서 그 동안 MRI 검진의 대상이 아니었던 치밀유방 여성에서의 적용을 위한 임상시험이 제안되었다 (10). 미국의 경우 검진 대상 연령군 (40-74세)의 약 43%가 치밀유방이며 우리나라와 동아시아 여성에서는 그 비율이 더 높다 (11, 12).

미국예방정책국특별위원회 (U.S. Preventive Services Task Force, USPSTF)는 평균 위험도의 치밀유방 여성에서 정상 유방촬영인 경우 초음파, DBT, MRI를 비롯한 검진은 추가적으로 암을 발견하지만 위양성도 증가시키며 유방암 결과에 대한 영향이나 유용성을 주장하기에는 과학적 근거가 불충분하다고 평가하였다 (13). 그 이유는 기존 연구들에서 보여주었던 발견 암 숫자의 단기적 증가만으로 추가 검진이 치료와 관련된 유병률이나 유방암 사망률을 줄이고 여성의 삶의 질을 향상시킨다고 주장할 수 없기 때문이다. 추가 검진의 유용성과 가치를 주장하기 위해서는 평균 위험도이며 치밀 유방인 여성에서 추가 검진을 받은 군과 받지 않은 군에서 진단시 유방암 병기, 재발률 및 사망률과 관련성이 높다고 알려진 간격암의 감소와 이에 따른 진행성 암 발생 감소 등의 지표를 위양성률과 함께 비교해야 한다 (14). 위양성은 암의심 병변으로 조직검사 필요로 판정했으나 양성으로 판정된 경우뿐만 아니라 양성 병변에 대해 비정상적으로 판정하여 추가 진단검사를 하거나 이후 단기 추적검사를 하는 경우를 포함하며 검진 참여 여성의 정신적 고통과 경제적 손실을 초래한다 (15). 최근에는 환자의 일생 동안 증상이나 죽음을 야기하지 않

본 종설은 2017년도 암정복추진연구개발사업의 연구비 지원을 받아 이루어진 것임 (HA17C0056)

통신저자: 문우경

(03080) 서울시 종로구 연건동 28번지

서울대학교 의과대학 영상의학교실

Tel. (02) 2072-3928, Fax. (02) 743-6358

E-mail: moonwk@snu.ac.kr

을 질환을 발견하는 과잉진단 (overdiagnosis)이란 개념이 암 검진의 위해로 등장하면서 새로운 검진 방법의 가치를 평가하는 일이 더 복잡해졌으며 (16, 17) 0기암 (관상피내암)을 제외한 침윤암 발견율과 종양 악성도에 따른 민감도 프로파일을 검진 방법간에 비교하는 것이 중요해졌다 (18, 19).

본 중설에서는 치밀유방 여성에서의 검진 유방촬영에 추가 검사로서 유방 MRI의 유효성 평가와 관련되어 진행된 DENSE와 EA1141 임상시험의 연구설계와 최근 NEJM과 JAMA에 발표된 중간 결과를 소개하고 그 의미 및 임상적용 방향에 대해 토의하고자 한다. DENSE와 EA1141 임상시험의 목적 및 계획과 MRI 검진 성적을 표 1과 2에 비교 정리했다 (20, 21).

I. DENSE 시험

1. 임상시험 목적

DENSE (Dense Tissue and Early Breast Neoplasm Screening) 시험의 목적은 극도 치밀유방 여성에서 유방촬영 단독 검진과 비교한 MRI 추가 검진의 간격암 발생 감소 효과를 평가하는 것이다 (20).

2. 임상시험 계획

(1) 시험 설계

네덜란드의 2년주기 인구기반 검진 프로그램 대상인 50-75세 여성 중에서 극도 치밀유방이며 유방촬영이 정상인 여성을 추가 MRI 검진에 초대된 군 (MRI 초대군)과 유방촬영 검진만 받는 군 (유방촬영 단독군)으로 1:4로 배

표 1. 두 임상시험의 목적과 계획 비교

	DENSE 시험	EA1141 시험
시험 목적	극도 치밀유방 여성에서 MRI 추가 검진의 간격암 발생 감소 효과 평가	치밀유방 여성에서 AB MRI와 DBT 검진 성적 비교
시험 설계	전향적 무작위임상시험	전향적, 개체 내, 비교 연구
대상군	MRI 초대군과 유방촬영 단독군으로 1:4 배정 네덜란드의 인구기반 검진 참가자 50-75세 (평균, 54세)	AB MRI과 DBT 검사와 판독을 무작위 순서로 배정 미국과 독일의 병원 검진 참가자 40-75세 (평균, 54세)
1차 평가변수	간격암률 차이	침윤암 발견율 차이
2차 평가변수	소환율, 암발견율, 위양성률, 양성예측도 및 민감도	민감도, 특이도, 추가 영상검사 권고율 및 생검의 양성예측도
통계와 표본크기	종양 특징 (병기, 등급, 수용체 상태) 삶의 질과 비용효과 비교 간격암률 차이, 1.95 (단측 검정, 유의수준 0.05, 검정력 80%) MRI 초대군 7,237명 (실제 검사자 4,776명)과 유방촬영 단독군 29,948명	종양 특징과 계층 프로파일 참가자 보고 결과 침윤암 발견율 차이, 9 (양측 검정, 유의수준 0.05, 검정력 90%) 1,450명

표 2. 두 임상시험의 MRI 성적 비교

	DENSE 시험	EA1141 시험
분석 대상자 수 (명)	MRI 검진 초대군 8,061	AB MRI와 DBT 모두 시행한 1,444
간격암률 (/1,000 검진)	2.5 (1.6-3.8)	0 (0-2.7)
전체 암발견율 (/1,000 검진)	16.5 (13.3-20.5)	15.2 (10.1-23.0)
침윤암 발견율 (/1,000 검진)	13.4 (10.5-17.1)	11.8 (7.4-18.8)
소환율 (또는 추가 영상 권고율) (%)	9.5 (8.7-10.4)	7.5 (6.2-9.0)
생검 양성예측도 (%)	26.3 (21.7-31.6)	19.6 (13.2-28.2)
민감도 (%)	95.2 (88.1-98.7)	95.7 (79.0-99.2)
특이도 (%)	92.0 (91.2-92.8)	86.7 (84.8-88.4)

괄호 안은 95% 신뢰구간.

정하여 3번의 검진을 시행하도록 설계되었다 (22). 2011년 12월 첫 대상자가 등록되었고 2015년 11월에 등록이 완료되었다.

(2) 대상군 선정

극도 치밀유방 판정에는 자동화 밀도측정 소프트웨어인 Volpara (Volpara Health Technology, Wellington, New Zealand)를 사용하였고 Volpara 유방밀도 4 (최신 버전 도입 후는 D)인 경우를 포함시켰다. 중증 신장 기능 장애 (사구체 여과율, 40 mL/분 미만), 가톨릭증 기반 조영제에 대한 이전의 부작용, 임신, 폐소 공포증 및 극심한 과체중 (150 kg 이상) 등 MRI 검사에 대한 표준 금기 사항이 있는 경우는 중재군에서 제외되었다.

(3) 영상검사 및 판독

국가 검진의 유방촬영은 모두 디지털 장비 (Selenia; Hologic, Bedford, Mass, USA)를 사용하였다. MRI는 3T 기기를 사용하였고 모든 환자에서 가도부트를 (Gadovist; Bayer Healthcare Medical Care, Berlin, Germany) 조영제를 사용했으며 통일된 프로토콜로 영상을 획득했다 (22). BI-RADS에 따라 판독하며 카테고리 4 이상의 병변은 조직검사를 시행했다. BI-RADS 카테고리 3인 참가자의 경우 MRI의 이중 판독이 수행되었고, 카테고리 3 판정에 대한 합의가 있는 경우 6개월 후 MRI를 사용한 후속 영상 검사를 시행했다. 후속 MRI의 결과는 음성 (정기적 검진 프로그램으로 복귀하는 BI-RADS 1 또는 2) 또는 양성 (BI-RADS 4 또는 5)으로만 보고토록 했다 (22).

(4) 평가변수

1차 평가변수 (primary end point)는 2년의 검진 기간 동안의 두 군간 간격암률 (interval-cancer rate) 차이로 간격암에는 정상 판정 후 다음 예정된 검진 전에 발견된 모든 유방암이 포함된다. 2차 평가변수 (secondary end point)는 추가 검사를 위한 소환율, 암발견율, 위양성률, 양성예측도 (positive predictive value, PPV)와 민감도 등 진단 성적, 종양 특징 (크기, TNM 병기, 등급, 수용체 상태와 분자아형) 및 삶의 질과 비용효과 비교이다. 두 군의 간격암 수와 종양 특징에 관한 데이터는 네덜란드 암 등록사업과 연계를 통해 수집했다. 이상 반응과 심각한 이상 반응은 MRI 검사 직후에 임상시험센터에서 기록하거나 참가 여성들에게 30일 이내에 보고토록 했다.

(5) 통계

MRI 초대군에서 실제 MRI 검사를 받는 참가율을 66%로 가정하여 1,000회 검진 당 간격암률 1.95의 군간 차이를 탐지하기 위해 필요한 시험 대상자 수를 계산했다. 단측 검정, $\alpha=0.05$ 유의 수준과 80%의 검정력을 갖기 위해 최소 MRI 초대군에 7,237명 (실제 검사자 4,776명)과 유방촬영 단독군에 29,948명의 시험 대상자 수가 필요했다. 두 군간 간격암률 차이는 χ^2 테스트를 사용하였고 데이터는 배정된 대로 분석 원칙 (intention-to-treat principle)과 순응평균인과효과 (complier average causal effect, CASE) 분석법을 적용했다 (23).

3. 첫 번째 검진의 결과

(1) 분석 대상군의 특징

유방촬영 검진을 받은 40,373명의 여성 중 8,061명 (20%)이 MRI 초대군에, 32,312명은 유방촬영 단독군에 할당되었다. MRI 초대군 8,061명의 여성 중 4,783명 (59%)이 실제로 MRI 검사를 받았다. MRI 초대군과 유방촬영 단독군은 나이, 거주 지역, 사회경제적 수준, 도시화 정도 등 기본 특성에 대해 균형을 이루었다.

(2) 1차 결과 지표

MRI 초대군 8,061명에서 20개의 간격암이 진단되었다. 20개 간격암 중 4개는 실제로 MRI를 검사한 여성 (검진 1,000회당 0.8)에서 진단되었고 16개는 초대를 수락하지 않은 여성 (검진 1,000회당 4.9)에서 진단되었다. 유방촬영 단독군 32,312명에서는 161개의 간격암이 진단되었다. 배정된 대로 분석 원칙에 따른 간격암률은 검진 1,000회당 MRI 초대군에서 2.5 (95% 신뢰구간 [confidence interval, CI], 1.6-3.8)와 유방촬영 단독군에서 5.0 (95% CI, 4.3-5.8)으로 검진 1,000회당 간격암률 차이는 2.5 (95% CI, 1.0-3.7)로 MRI 초대군에서 낮았다 ($P<0.001$). 관상피내암으로 진단된 여성을 제외해도 결과는 변하지 않았다.

대조군인 유방촬영만 받은 여성 중 MRI 검진을 수용했을 여성과 수용하지 않았을 여성을 각각 59%와 41%로 가정한 CASE 분석에 의하면 (23) 검진 1,000회당 간격암률 차이는 4.2 (95% CI, 2.0-6.4)로 MRI 검진을 수용할 여성에서 간격암률은 유방촬영만 받은 여성에서보다 낮았다 ($P<0.001$).

(3) 2차 결과 지표

실제 MRI를 시행한 4,783명 중 453명이 이상판정을 받아서 소환율은 검진 1,000회당 94.9 (95% CI, 86.9-103.6)였다. 이 중 300명이 조직검사를 받아 생검률은 검진 1,000회당 62.7 (95% CI, 56.2-70.0)이었고 이 중 79명 (64 침윤암과 15 관상피내암)에서 암이 발견되어 검진 1,000회당 암발견율은 16.5 (95% CI, 13.3-20.5)였고 침윤암 발견율은 13.4 (95% CI, 10.5-17.1)였다. 소환검사의 양성예측도 (PPV1)는 17.4% (95% CI, 14.2-21.2), 생검의 양성예측도 (PPV3)는 26.3% (95% CI, 21.7-31.6)였고 위양성은 검진 1,000회당 79.8 (95% CI, 72.4-87.9), 특이도는 92.0% (95% CI, 91.2-92.8)였다.

MRI 검진의 민감도는 95.2% (95% CI, 88.1-98.7)였다. 검진에서 발견된 종양 크기는 다른 그룹의 사람들과 비교해서 평균적으로 더 작았다. 침윤암의 중간 크기는 MRI 참가자 중 검진으로 발견된 경우는 9.5 mm (사분범위, 6.8-12.0), MRI 참가자 중 간격암으로 발견된 경우는 13.0 mm (사분범위, 10.5-17.0), MRI 비 참가자의 경우는 15.0 mm (사분범위, 12.0-31.0), 유방촬영 단독군 여성에서는 17.0 mm (사분범위, 12.0-23.0)였다.

MRI 참가자 중에서, 침윤암 (유관암 및 소엽암)의 절대 발생률은 유방촬영 단독군보다 높았으며 병이 없는 질병 과정을 겪을 수도 있는 관상피내암 및 관상암 (tubular cancers)의 절대 발생률도 높았다. 관상피내암은 MRI 참가자에서 0.31% (15/4,783)인 반면 유방촬영 단독군에서는 0.03% (9/32,312)였다. MRI 참가자들에서 림프절 음성 및 초기 암의 절대 발생률도 높았다. 이 소견은 약간 낮은 후기 암의 발견율과 동반되었지만, 이러한 암의 숫자는 적었다. MRI 초대군에서 검출된 암은 유방촬영 단독군에서보다 더 잘 분화되고 호르몬-수용체 양성인 것으로 나타났다. MRI 참가자에서 발견되고 등급 정보가 있었던 침윤암의 종양 등급은 저등급 31개, 중간등급 24개와 고등급 4개인 반면 간격암은 저등급 0개, 중간등급 2개와 고등급 1개였고 MRI 비참가자 중 간격암은 저등급 0개, 중간등급 7개와 고등급 4개였으며 유방촬영 단독군에서 간격암은 저등급 29개, 중간등급 70개와 고등급 39개였다. 다음 번 유방촬영 검진에서, 암발견율은 MRI 참가자 중 검진 1,000회당 2.0인 반면 MRI 비참가자는 7.1, 유방촬영 단독군은 6.0이었다. 이러한 결과는 MRI 검사가 암의 진단 시간을 앞당겼음을 의미한다.

(4) 이상 반응

MRI 참가자 중 0.1%는 검사 중 또는 직후에 이상 반

응 또는 심각한 이상 반응을 나타냈으며 혈관 반응, 조영제에 대한 반응 또는 정맥 내 주입과정과 관련 된 것들이었다. MRI 검사와의 관련성 여부와 상관없이 모든 건강 문제를 조사한 30일 설문지에서 MRI 참가자의 25.7% (1,233/4,783)에서 근골격계 이상 및 호흡곤란 등 이상 반응을 보고했고 0.6% (27/4,783)에서는 응급실 또는 병원 방문이 필요한 심각한 이상 반응을 보고했다.

II. EA1141 시험

1. 임상시험 목적

EA1141 시험의 목적은 유방암 위험이 평균이며 치밀유방 조직을 가진 여성에서 AB MRI와 DBT 검진의 성적을 참가자 수준에서 비교하는 것이다 (21).

2. 임상시험 계획

(1) 시험 설계

미국의 47개 기관과 독일의 1개 기관이 참여한 전향적, 개체 내, 비교, 종단 추적을 이용한 단면 연구 (cross-sectional study with longitudinal follow-up)로 ECOG-ACRIN 암 연구 그룹 (<https://ecog-acrin.org>)에 의해 진행되었다. DBT로 정기 유방암 검진을 받을 비균질 또는 극도 치밀 유방 여성에서 연구 기준 시점 및 1년 후에 DBT 및 AB MRI로 영상 검사를 받고 3년간 추적 관찰을 계획하였다. 중앙 무작위화 방법 (대체 없이 크기 4의 치환된 블록을 사용하는 1:1 비율 방식)을 사용하여 영상 검사의 순서를 결정했다.

추적 조사를 위해, 처음 검사 후 및 2차 검진 라운드 전 6개월 (±1개월) 및 11-13개월에 전화로 여성들에게 연락하였고 처음 검사 이후 유방 영상을 받았는지 여부, 유방암 진단을 받았는지 여부 그리고 유방암이 진단되었다면 발견 방법을 여성들에게 질문했다. 코어 또는 외과적 생검의 병리 결과가 암발견율 및 생검의 양성 예측도에 대한 기준으로 사용되었고 11-13개월의 추적 관찰 기간 동안 (다음 검진 전까지) 보고된 간격암 역시 민감도와 특이도에 대한 기준에 포함되었다. 2016년 12월에 첫 대상자가 등록되었고 2017년 11월에 등록이 완료되었다.

(2) 대상군 선정

40-75세 무증상 여성으로 최근 검진 유방촬영에서 치밀 유방으로 보고된 경우에 등록했다. 양성 유방 생검 병력, 치료 유방암 병력 또는 유방암 가족력이 있는 여성도

포함되었다. 지난 12개월 내에 유방 초음파, 유방 MRI, 분자 유방 영상 또는 조영증강 유방촬영 검사를 받았거나 미국 암 협회 지침에 따라 전체 프로토콜 유방 MRI를 받을 자격이 있는 여성은 제외되었다 (24). 유방암의 개인 병력이 없고 75세 미만인 참가자는 유방암 감시 컨소시엄 (Breast Cancer Surveillance Consortium, BCSC) 위험 계산기로 개별 위험 점수를 평가했다 (25). BCSC 5년 위험 점수는 저위험 (<1.0%), 평균위험 (1.0-1.66%), 중간위험 (1.67-6.0%) 및 고위험 (>6.0%)으로 구분된다.

(3) 영상검사와 판독

모든 AB MRI 연구는 10분 미만의 총 획득 시간으로 검사가 완료해야 했으며, 가도베네이트 디메글루민 (MultiHance; Bracco Diagnostics Inc, Monroe Twp., NJ, USA) 주입 전과 후에 T2 및 T1 가중 영상 획득을 포함했다. 모든 연구는 유방 영상보고 및 데이터 시스템 (BI-RADS)에 따라 카테고리 1-5로 판독되었다. 판독에 참가할 영상의학과 전문의들은 AB MRI 판독 교육 후에 증례 시험에서 합격해야 자격이 주어졌다. 영상 검사는 다른 검사의 결과를 모르게 한 2명의 유방 영상의학과 전문의에 의해 독립적으로 해석되었다. DBT와 AB MRI 모두 24시간 내에 수행되었다. 검진 유방촬영 또는 DBT에서 양성소견이 발견되면, 여성은 추가 영상 워크업을 위해 재방문을 했지만 AB MRI의 경우는 추가 영상 검사 없이 바로 최종 판정을 실시했다. 두 검사 모두에서 같은 쪽 유방에 양성소견이 있는 경우는 영상의학과 의사가 동일한 병변인지 다른 병변인지 판정했다.

(4) 평가변수

1차 평가변수는 처음 검진에서 각 영상검사에 의해 발견되고 병리학(코어 생검 또는 외과 절제)적으로 확인된 참가자 수준에서의 침윤암 발견율이다. 2차 평가변수는 민감도, 특이도, 추가 영상검사 권고율 (즉, 재검 및 단기 추적 권고) 및 생검의 양성예측도 (PPV3)이다. 침윤암 및 관상피내암의 존재가 평가의 양성 기준 표준으로 정의되었다. 연구의 추가 보조 평가 항목은 참가자 보고 결과 (participant-reported outcomes), 게놈 프로파일에 의한 종양 생물학, 양성예측도 및 2차 검진 라운드의 추가 영상 권고율 및 3년간 암발견율로 구성된다. 추가적인 탐색 평가 변수에는 전체 암발견율 (침윤암 및 관상피내암), 간격암률, 병변 수준의 양성예측도 추정치, 진단시 암의 병기 및 조직 병리학과 면역조직화학적 특징이다.

(5) 통계

표본 크기는 양측 검정, $\alpha=0.05$ 유의 수준에서 두 검사 간에 1,000명당 9명의 침윤암 발견율의 차이를 90% 검정력으로 제공하도록 계산되었으며 1,450명의 시험 대상자 수가 필요했다. 예상 프로토콜 개발 당시 평균 위험 여성에서 AB MRI 검진 대 DBT에 대한 연구 발표가 없었기 때문에 표본 크기 추정이 어려웠다. 따라서, 분석은 디지털 유방촬영과 비교하여 1,000명당 6-11의 추가적인 침윤암 발견율을 보고한 평균 위험 여성의 MRI 검진 연구의 제한된 증거를 사용했다 (12, 26). 쌍으로 된 디자인으로 exact McNemar 테스트를 사용하여 비율 (발견율, 민감도, 특이도 및 추가 영상 권고율)을 비교했다. 침윤암 발견율 차이는 Bonett-Price Laplace 조정을 사용한 Wald 간격으로 계산했다. PPV3 (참가자 수준 및 병변 수준 둘 다)의 비교는 일반화된 추정 방정식 (generalized estimating equation, GEE) 회귀 모델링을 사용하여 수행되었다.

3. 첫 번째 검진의 결과

(1) 분석 대상군의 특징

1,516명 등록자 중 두 검사를 모두 시행한 1,444명 (평균연령 54세; 40-75세)이 분석 대상에 포함되었다. 이 여성들의 DBT에서 유방 밀도는 8%가 지방 또는 유선산재였고 77%는 비균질 치밀, 15%는 극도 치밀이었다. BCSC에 의한 5년 위험 점수의 중앙값은 1.6% (범위, 0.3-7.8), 평균은 1.8% (표준편차, 0.8)였다. 12명의 참가자는 74세 이상 ($n=4$)이거나 유방암 진단을 받은 사람 ($n=8$)이므로 위험 점수를 계산할 수 없었다. 또한, 유방암 병력이 불확실한 47명의 참가자는 위험 점수 계산에서 제외되었다.

(2) 1차 결과 지표

AB MRI 검사는 이 기간 동안 발생한 23개 유방암 중에서 17개 침윤암 모두와 관상피내암 6개 중 5개를 발견하였다. DBT는 17개 침윤암 중에서 7개와 관상피내암 6개 중에서 2개를 발견하였다. 침윤암 발견율은 AB MRI가 1,000명당 11.8 (95% CI, 7.4-18.8)이었고 DBT가 4.8 (95% CI, 2.4-10.0)로 그 차이는 1,000명당 7 (95% CI, 2.2-11.6) ($P=0.002$)이었다.

(3) 2차 결과 지표

침윤암과 관상피내암 발견에 대한 AB MRI와 DBT의 민감도는 각각 95.7% (95% CI, 79.0-99.2)와 39.1%

(95% CI, 22.2-59.2)로 유의한 차이가 있었다 ($P=0.001$). 하지만, AB MRI의 특이도는 86.7% (95% CI, 84.8-88.4)로 DBT의 97.4% (95% CI, 96.5-98.1) 보다 유의하게 낮았다 ($P<0.001$). 6개월 추적검사가 필요했던 카테고리 3 판정은 AB MRI에서 7.5% (95% CI, 6.2-9.0; 108/1,444), DBT의 경우는 1.2% (95% CI, 0.8-2.0; 18/1,444)였다. 추가 영상 권고율은 AB MRI의 경우 7.5% (95% CI, 6.2-9.0)였고 DBT의 경우는 10.1% (95% CI, 8.7-11.8) ($P=0.02$)였으며 생검의 양성예측도는 각각 19.6% (95% CI, 13.2-28.2)와 31.0% (95% CI, 17.0-49.7) ($P=0.15$)였다. 병변 수준에서 생검의 양성예측도는 각각 19.0% (95% CI, 12.6-27.7)와 35.5% (95% CI, 19.5-55.5) ($P=0.08$)였다. 참가 기관 중에는 AB MRI를 임상에서 사용하지 않는 병원도 있었지만 시험에 참가한 기관들의 검진 성적은 비슷하였다.

전체 암발견율은 1,444명 중에 23명으로 여성 1,000명당 15.9이었다. 23명의 여성 중 AB MRI는 22명에서, DBT는 9명의 여성에서 암을 발견해서 1,000명당 암발견율은 AB MRI가 15.2 (95% CI, 10.1-23.0), DBT가 6.2 (95% CI, 3.3-11.8) ($P=0.001$)였다. 간격암은 추적기간 동안 보고되지 않아서 간격암률은 0% (95% CI, 0-2.7)였다. 침윤암의 중앙 크기는 10.5 mm (범위, 4-48 mm)였고 림프절 음성률은 96% (22/23)였다. DBT와 AB MRI 모두에서 보였던 7명의 침윤암 중앙 등급은 저등급 3개와 중간등급 4개였다. AB MRI에서만 보였던 10명의 침윤암 중앙 등급은 저등급 3개, 중간등급 4개와 고등급 3개였다.

(4) 이상 반응

등록 후 1년 이내에 12명 (0.8%) 여성 중 총 13건의 이상 반응이 보고되었다. 8건 (62%)은 1등급 이하였고 5건 (38%)은 2등급 이상이었고 이 중 한 명은 뇌출혈로 사망했다. MRI 검사와 관련된 가장 흔한 부작용은 가벼운 알레르기 반응 (3건)과 불안 (2건)이었다.

III. 시험 결과의 해석과 토론

1. 참가자 특징

(1) 유방밀도 비교

기존 MRI 연구들은 BRCA 유전자 변이 보인자이거나 가족력이나 유방암 과거력 등을 기반으로 한 유방암 위험도 평가결과에서 중간 또는 고위험군 여성을 대상으로 했지만 DENSE나 EA1141 임상시험에서는 유방밀도를 기준

으로 MRI 검진 여성을 선별했다. DENSE 시험에서는 국가 검진에 참가한 여성 중에서 극도 치밀유방의 여성만을 선별했고 전체의 20%가 Volpara로 측정했을 때 극도 치밀유방을 가진 여성이었다. 네덜란드 국가 검진 자료에 의하면 극도 치밀유방의 여성은 치밀하지 않은 유방이나 비균질 치밀유방의 여성과 비교해서 유의하게 민감도가 떨어지고 간격암률이 높았다 (27). 반면 EA1141 시험에서는 48개 대학 또는 지역 의료기관에서 이전 유방촬영에서 치밀유방 즉 BI-RADS C 또는 D로 판독된 여성을 등록했고 DBT 검사에서 참가자의 실제 유방밀도는 8%가 지방 또는 유선산재였으며 92%만 치밀유방 (비균질 치밀 77%와 극도 치밀 15%)이었다. EA1141 시험에서 비균질 치밀유방과 극도 치밀유방의 여성간에 유방밀도에 따른 암발견율 차이는 없었다.

(2) 유방암 위험도

DENSE 시험에서, 대조 군의 오염을 방지하기 위해 Zelen의 사전 무작위 배정 설계가 사용되었고 결과적으로 중재 군과 대조 군의 모든 비 참가자는 사전 동의를 제공하지 않았다 (28). 따라서 이들 여성들에 대한 정보는 우편번호 만으로 참가자와 비 참가자 간의 유방암 위험도에 대한 비교는 제한적이다. DENSE 임상시험 참가자 중 유방암 발견 여성에 대한 위험도 조사가 향후 진행될 예정이다 (van Gils CH, personal communication).

EA1141 시험에서는 참가자 개인에 대해서 BCSC 모델에 의한 5년과 10년 위험 점수를 계산하였다. 연구 참가자의 BCSC 모델에 의한 5년 위험 점수의 평균은 1.8%로 절반 정도가 중간위험 (1.67-6.0%) 및 고위험 (>6.0%) 여성이었다. 또한, 침윤암 진단 여성 중 BCSC 위험 점수 정보를 가진 16명 여성 중에서 15명 (94%)이 중간위험 또는 고위험 여성에서 발견되었다. 이상은 치밀유방 여성에서 추가 검진 효율성을 높이기 위해서는 위험도 평가가 중요함을 시사한다 (29). BCSC 모델을 사용한 대규모 연구에서 5년 위험이 1.67% 이상이고 극도 치밀유방 또는 5년 위험이 2.50% 이상이고 비균질 치밀유방인 여성에서 높은 간격암 발생률이 관찰되었다 (30, 31). 유방 밀도를 포함한 모델이 기존 유방암 위험 요인만으로 계산한 위험 평가 모델보다 예측력이 높은 것으로 보고되었으며 (32, 33) BCSC와 Tyrer-Cuzick 모델에는 BI-RADS나 Volpara 유방밀도가 포함되어 있다.

2. 결과의 해석

(1) 간격암, 침윤암 발견율과 발견된 암의 특징

DENSE 시험은 MRI 초대군과 유방촬영 단독군에서 검진 시행 후 간격암률을 비교할 수 있게 설계되었다. 그 결과는 유방 조직이 극도로 치밀한 여성에서 MRI 추가 검진을 하면 유방촬영 단독으로 검진을 하는 경우보다 2년 검진 기간 동안 간격암 발생이 절반으로 줄어 든다는 것이다. MRI 참가자들에서 림프절 음성 및 초기 암의 절대 발생률도 높았다. 하지만 림프절 양성 암의 빈도는 검진 1,000회당 실제 MRI 검사자와 유방촬영 단독군에서 각각 1.9와 2.2로 비슷했다. 간격암 감소가 진행성 암의 감소와 사망률 감소로 연결될 것인가가 가장 중요한 이슈이다. 간격암 감소가 진행성 암의 감소에 영향을 미치는지 보려면 최소 4-5년의 추적기간이 필요하다 (34, 35). 사망률 감소 효과는 Microsimulation Screening Analysis (MISCAN) 컴퓨터 모델로 추후 계산할 예정이다 (22).

EA1141 시험에서는 DBT와 AB MRI의 검사와 판독의 순서를 무작위로 정한 후 두 가지 검사를 모두 시행한 여성에서 1,000명당 침윤암의 차이를 비교할 수 있게 설계되었다. 그 결과는 침윤암 발견율은 AB MRI가 1,000명당 11.8이었고 DBT가 4.8로 그 차이는 1,000명당 7이었다. DBT와 AB MRI 모두에서 보였던 7명의 침윤암에는 고등급 암이 없었지만 AB MRI에서만 보였던 10명의 침윤암 중에서 3명은 고등급 암이었으며 9명은 호르몬-수용체 양성 암이었다. DBT와 AB MRI에서 발견된 침윤암의 생물학적 특성을 알기 위해 기존의 종양 조직학적 등급이나 수용체 상태뿐만 아니라 PAM50 에세이와 Oncotype DX 재발 점수 등 유전체 수준의 분석을 진행하고 있다.

하지만 과잉진단과 관계가 있는 관상피내암은 MRI 참가자들에서는 0.31%인 반면 유방촬영 단독 군에서는 0.03%였다. 두 임상 시험 모두 MRI에서 추가 발견 암의 과반수 이상이 관상피내암 또는 호르몬-수용체 양성인 저등급 침윤암인 점을 고려하면 과잉진단의 문제는 상당부분 남아 있다. 검진 MRI 임상 시험의 결과가 참여자의 사망률 감소뿐만 아니라 치료와 관련된 유병률을 줄이고 여성의 삶의 질을 향상시킬 수 있는 지에 대한 추적 데이터가 필요하다.

(2) 위양성률 비교

두 연구 모두에서 전체 프로토콜 또는 AB MRI 검사는 디지털 유방촬영이나 DBT에 비해 위양성률이 높았다. DENSE 시험에서는 MRI 검사 이후에 10%정도가 재검

이 필요했고 300명이 MRI에서 암 의심 소견으로 생검을 하게 되었으며 수행되는 생검의 74%가 암 진단으로 이어지지 않았다. EA1141 시험에서 생검이 필요한 경우가 AB MRI에서는 107명이었고 DBT에서는 29명으로 생검의 양성예측도는 각각 19.0%와 35.5%였다. 6개월 추적검사가 필요했던 카테고리 3 판정은 AB MRI와 DBT에서 각각 7.5%와 1.2%였다. 하지만 MRI 검진의 위양성률은 최초 검진인 경우가 가장 높고 이전 비교 가능한 MRI 영상이 있는 후속 검진들에서는 보통 절반 수준으로 줄어든다 (36). MRI 검사의 위양성률은 판독의사의 전문성과 판독의 양에 좌우되므로 교육과 학습의 역할이 중요하다. 확산강조영상이나 인공지능 (artificial intelligence, AI)-컴퓨터보조진단 (computer-aided diagnosis, CAD)의 사용도 MRI의 위양성률을 줄일 수 있다 (37, 38).

3. 안전성과 편의성

(1) 참가자의 이상 반응 보고

DENSE 시험 참가자 중 0.1%와 EA1141 참가자 중 0.8%는 MRI 검사 중 또는 직후에 혈관 반응, 조영제에 대한 반응 또는 정맥 내 주입과정과 관련 된 이상 반응 또는 심각한 이상 반응을 나타냈다 (20, 21). 신기능이 심하게 저하된 사람이나 임산부 또는 이전에 MRI 조영제에 과민 반응 병력이 있는 환자에서는 가돌리늄 조영제를 사용할 수 없다. 조영제를 사용하는 AB MRI도 이러한 문제점들을 그대로 가지고 있다 (39). 최근 조영제를 사용하지 않고 신속 영상이 가능하며 고등급 침윤암에 대한 민감도가 높은 DWI 프로토콜을 이용한 검진에 대한 관심이 증가하고 있는 이유이다 (40).

(2) (비)참가자에 대한 설문조사

DENSE 시험에서 중재 군에 무작위 배정 된 여성 중 실제 MRI 검사를 받은 사람은 59%였다. 비 참여에 대해 가장 자주 언급 된 이유로는 “MRI 관련 불편” (예: 폐소 공포증, 조영제 거부), “MRI 금기 사항” (27%) 및 “보충 검사 결과에 대한 불안” (21%)이었다 (41). 참여에 관심이 있는 여성은 참여의 이유로 “예상되는 개인 건강 혜택” (68%)과 “과학에의 기여” (43%)를 가장 자주 언급했다. “추가 검사 결과에 대한 불안”은 비 참여에 중요한 이유이므로, 같은 날에 두 가지 기술을 제공하고, 하나의 종합 평가 결과를 제공하며, 명확하고 증거 기반의 혜택 및 피해에 대한 정보를 제공하면 장벽을 더욱 줄일 수 있을 것이라고 저자들은 제안하였다. ACRIN 6666 연구에서 MRI

검사 참가율은 66%였으며 초대를 거절한 사람들의 사유를 비교하면 DENSE 시험에서 조영제 사용에 대한 거부 의 비중이 높은 편이다 (42). 이것은 최근의 가돌리늄 뇌 침착 등 안정성에 대한 이슈를 반영한 것으로 평가된다.

(3) 비용과 접근도

높은 비용이 유방 MRI 검사의 검진 사용에 주요한 장애 요인이므로 AB MRI는 검사 및 판독 시간을 단축시켜서 검사비를 \$300-\$500 (미국 기준) 정도로 만들어 MRI 검진의 문턱을 낮추고자 한다. 가치 기반의 의료에서 가치는 비용에 대한 환자의 결과로 정의되므로, AB MRI는 비용효과를 높여 MRI 검진의 가치를 향상시킬 수 있을 것이다 (43). 하지만, 검진 MRI가 이득이 확실한 고위험군에서 조차 보편화되어 있지 않은 상황에서 개인이 비용을 지불해야 하는 AB MRI의 확산은 유방암 검진 수혜의 불균형을 높일 수 있다는 우려도 있다 (39, 44).

결 론

표준 검진 방법인 유방촬영은 생물학적으로 공격적인 암에 대한 민감도가 낮은 과소진단의 문제와 일생 동안 증상이나 죽음을 야기하지 않을 질환을 상대적으로 많이 발견하는 과잉진단의 문제를 갖고 있다. 최근 수행된 DENSE와 EA1141 시험에서는 평균 위험도의 치밀유방 여성에서 MRI 추가 검진이 디지털 유방촬영이나 DBT 단독 검진과 비교해서 진행성 유방암 감소의 대리 지표인 간격암의 감소와 침윤암의 추가 발견에 유용한 지를 평가했으며 그 결과는 MRI 추가 검진이 간격암 감소와 침윤암 발견에 효과가 있음을 보여주었다. 하지만 추가 발견 암의 과반수 이상이 관상피내암 또는 저등급 침윤암인 점을 고려하면 과잉진단의 문제는 상당부분 남아 있어서 두 임상 시험의 결과가 실제 검진에 수용되기 위해서는 MRI 검진이 유방암 진단 후 치료와 관련된 유병률을 줄이고 사망률 감소로 이어질 수 있는지에 대한 장기 추적 데이터가 필요하다. 전체 생존율 개선에 기여하는지 불분명한 상태에서 고가의 MRI 검사를 평균 위험도의 무증상 여성의 검진에 도입하는 것은 의료비 증가뿐 아니라 상당수의 여성에서 위양성 증가와 조영제 부작용 등에 의한 피해를 초래할 것이다. 그러므로 치밀유방 여성 중에서 추가 MRI 검진이 필요한 고위험군을 선별하기 위한 도구 및 위양성률과 비용을 더욱 줄일 수 있는 AB MRI 프로토콜 개발이 향후 중요한 연구 과제이다.

참 고 문 헌

1. 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
2. Park B, Cho HM, Lee EH, et al. Does breast density measured through population-based screening independently increase breast cancer risk in Asian females? *Clin Epidemiol* 2017;10:61-70
3. Keating NL, Pace LE. New Federal Requirements to Inform Patients About Breast Density: Will They Help Patients? *JAMA* 2019;321(23):2275-2276
4. <https://densebreast-info.org/about.aspx> (accessed March 24, 2020)
5. Sprague BL, Stout NK, Schechter C, et al. Benefits, harms, and cost-effectiveness of supplemental ultrasonography screening for women with dense breasts. *Ann Intern Med* 2015;162(3):157-166
6. Rafferty EA, Durand MA, Conant EF, et al. Breast Cancer Screening Using Tomosynthesis and Digital Mammography in Dense and Nondense Breasts. *JAMA* 2016;315(16):1784-1786
7. Berg WA, Zhang Z, Lehrer D, et al. Detection of breast cancer with addition of annual screening ultrasound or a single screening MRI to mammography in women with elevated breast cancer risk. *JAMA* 2012;307(13):1394-1404
8. Mann RM, Balleyguier C, Baltzer PA, et al. Breast MRI: EUSOBI recommendations for women's information. *Eur Radiol* 2015;25(12):3669-3678
9. Wernli KJ, DeMartini WB, Ichikawa L, et al. Patterns of breast magnetic resonance imaging use in community practice. *JAMA Intern Med* 2014;174(1):125-132
10. Kuhl CK, Schrading S, Strobel K, Schild HH, Hilgers RD, Bieling HB. Abbreviated breast magnetic resonance imaging (MRI): first postcontrast subtracted images and maximum-intensity projection—a novel approach to breast cancer screening with MRI. *J Clin Oncol* 2014;32(22):2304-2310
11. Sprague BL, Gangnon RE, Burt V, et al. Prevalence of mammographically dense breasts in the United States. *J Natl Cancer Inst*. 2014;106(10). pii: dju255.
12. Hong S, Song SY, Park B, et al. Effect of Digital Mammography for Breast Cancer Screening: A Comparative Study of More than 8 Million Korean Women. *Radiology* 2020;294(2):247-255

13. Melnikow J, Fenton JJ, Whitlock EP, et al. Supplemental Screening for Breast Cancer in Women With Dense Breasts: A Systematic Review for the U.S. Preventive Services Task Force. *Ann Intern Med* 2016;164(4):268-278
14. Irwig L, Houssami N, Armstrong B, Glasziou P. Evaluating new screening tests for breast cancer. *BMJ* 2006;332(7543):678-679
15. Myers ER, Moorman P, Gierisch JM, et al. Benefits and Harms of Breast Cancer Screening: A Systematic Review. *JAMA* 2015;314(15):1615-1634
16. Duffy SW, Agbaje O, Tabar L, et al. Overdiagnosis and overtreatment of breast cancer: estimates of overdiagnosis from two trials of mammographic screening for breast cancer. *Breast Cancer Res* 2005;7(6):258-265
17. Houssami N. Overdiagnosis of breast cancer in population screening: does it make breast screening worthless? *Cancer Biol Med* 2017;14(1):1-8
18. Kuhl CK. Abbreviated Magnetic Resonance Imaging (MRI) for Breast Cancer Screening: Rationale, Concept, and Transfer to Clinical Practice. *Annu Rev Med* 2019;70:501-519
19. Sung JS, Stamler S, Brooks J, et al. Breast Cancers Detected at Screening MR Imaging and Mammography in Patients at High Risk: Method of Detection Reflects Tumor Histopathologic Results. *Radiology* 2016;280(3):716-722
20. Bakker MF, de Lange SV, Pijnappel RM, et al. Supplemental MRI Screening for Women with Extremely Dense Breast Tissue. *N Engl J Med* 2019;381(22):2091-2102
21. Comstock CE, Gatsonis C, Newstead GM, et al. Comparison of Abbreviated Breast MRI vs Digital Breast Tomosynthesis for Breast Cancer Detection Among Women With Dense Breasts Undergoing Screening. *JAMA* 2020;323(8):746-756
22. Emaus MJ, Bakker MF, Peeters PH, et al. MR Imaging as an Additional Screening Modality for the Detection of Breast Cancer in Women Aged 50-75 Years with Extremely Dense Breasts: The DENSE Trial Study Design. *Radiology* 2015;277(2):527-537
23. Cuzick J, Edwards R, Segnan N. Adjusting for non-compliance and contamination in randomized clinical trials. *Stat Med* 1997;16:1017-1029
24. Saslow D, Boetes C, Burke W, et al. American Cancer Society guidelines for breast screening with MRI as an adjunct to mammography. *CA Cancer J Clin* 2007;57(2):75-89
25. <https://tools.bcscc-scc.org/BC5yearRisk/intro.htm> (accessed March 24, 2020)
26. Kuhl CK, Strobel K, Bieling H, Leutner C, Schild HH, Schrading S. Supplemental Breast MR Imaging Screening of Women with Average Risk of Breast Cancer. *Radiology* 2017;283(2):361-370
27. Wanders JOP, Holland K, Karssemeijer N, et al. The effect of volumetric breast density on the risk of screen-detected and interval breast cancers: a cohort study. *Breast Cancer Res* 2017;19(1):67
28. Zelen M. A new design for randomized clinical trials. *N Engl J Med* 1979;300(22):1242-1245
29. Longo DL. Detecting Breast Cancer in Women with Dense Breasts. *N Engl J Med* 2019; 381(22):2169-2170
30. Kerlikowske K, Zhu W, Tosteson AN, et al. Identifying women with dense breasts at high risk for interval cancer: a cohort study. *Ann Intern Med* 2015;162(10):673-681
31. Kerlikowske K, Sprague BL, Tosteson ANA, et al. Strategies to Identify Women at High Risk of Advanced Breast Cancer During Routine Screening for Discussion of Supplemental Imaging. *JAMA Intern Med*. 2019 Jul 1
32. Brentnall AR, Cuzick J, Buist DSM, Bowles EJA. Long-term Accuracy of Breast Cancer Risk Assessment Combining Classic Risk Factors and Breast Density. *JAMA Oncol* 2018;4(9):e180174
33. Dembrower K, Liu Y, Azizpour H, et al. Comparison of a Deep Learning Risk Score and Standard Mammographic Density Score for Breast Cancer Risk Prediction. *Radiology* 2020;294(2):265-272
34. Foca F, Mancini S, Bucci L, et al. Decreasing incidence of late-stage breast cancer after the introduction of organized mammography screening in Italy. *Cancer* 2013;119:2022-2028
35. Tabar L, Yen AM, Wu WY, et al. Insights from the breast cancer screening trials: how screening affects the natural history of breast cancer and implications for evaluating service screening programs. *Breast J* 2015;21(1):13-20
36. Bakker MF. MRI in addition to mammography screening in women with extremely dense breasts: outcome of the second (incident) round of the randomised DENSE trial [CTiR 11-11] ECR 2020

37. Pinker K, Moy L, Sutton EJ, et al. Diffusion-Weighted Imaging With Apparent Diffusion Coefficient Mapping for Breast Cancer Detection as a Stand-Alone Parameter: Comparison With Dynamic Contrast-Enhanced and Multiparametric Magnetic Resonance Imaging. *Invest Radiol* 2018;53(10):587-595
38. Verburg E, van Gils CH, Bakker MF, et al. Computer-Aided Diagnosis in Multiparametric Magnetic Resonance Imaging Screening of Women With Extremely Dense Breasts to Reduce False-Positive Diagnoses. *Invest Radiol* 2020 Mar 6
39. Tosteson ANA. An Abbreviated MRI Protocol for Breast Cancer Screening in Women With Dense Breasts: Promising Results, but Further Evaluation Required Prior to Widespread Implementation. *JAMA* 2020;323(8):719-721
40. Amornsiripanitch N, Bickelhaupt S, Shin HJ, et al. Diffusion-weighted MRI for Unenhanced Breast Cancer Screening. *Radiology* 2019;293(3):504-520
41. de Lange SV, Bakker MF, Monninkhof EM, et al. Reasons for (non)participation in supplemental population-based MRI breast screening for women with extremely dense breasts. *Clin Radiol* 2018;73(8):759.e1-759.e9
42. Berg WA, Blume JD, Adams AM, et al. Reasons women at elevated risk of breast cancer refuse breast MR imaging screening: ACRIN 6666. *Radiology* 2010;254(1):79-87
43. Leithner D, Moy L, Morris EA, Marino MA, Helbich TH, Pinker K. Abbreviated MRI of the Breast: Does It Provide Value? *J Magn Reson Imaging* 2019;49(7):e85-e100
44. Haas JS, Hill DA, Wellman RD, Hubbard RA, et al. Disparities in the use of screening magnetic resonance imaging of the breast in community practice by race, ethnicity, and socioeconomic status. *Cancer* 2016;122(4):611-617