


A black square containing the white text 'AI'.

AI

의료기술의 다면적 가치평가 예시집

2024. 5



본 예시집은 2023년 정부(보건복지부)의 재원으로 한국보건 의료연구원에서 수행한 연구사업인 「AI 소프트웨어 의료기기의 의료현장 적용을 위한 평가방안 마련 연구(NA23-002)」(주관연구책임자: 서울아산병원 영상학과 박성호 / 공동연구책임자: 한국보건 의료연구원 보건 의료연구본부 설아람)의 일환으로 제작되었습니다. 더 상세한 내용은 「한국보건 의료연구원 홈페이지 (www.neca.re.kr) - 연구정보 - 연구보고서」에서 확인하실 수 있습니다.

AI 의료기술의 다면적 가치평가 예시집

CONTENTS

	예시집 사용 관련 안내 사항	4
예시 1	현저한/의미 있는 진단정확도 향상, 진료/환자 결과 향상	6
예시 2	의료기관/의료인 효율 향상	12
예시 3	의료기관/의료인 효율 향상, 환자중심결과 향상	15
예시 4	거시적 의료비용 감소	19
예시 5	진료/환자 결과 향상	23
예시 6	진료/환자 결과에 대한 불분명한 효과, 의료기관/의료인 효율 감소	27
예시 7	의료기관/의료인 효율 향상	31
예시 8	진단정확도 향상에 대한 불분명한 효과, 의료기관/의료인 효율 향상에 대한 불분명한 효과	35
예시 9	진단정확도 향상에 대한 불분명한 효과, 의료기관/의료인 효율 감소	44
예시 10	진단정확도 향상에 대한 불분명한 효과, 진료/환자 결과에 대한 불분명한 효과	50
예시 11	환자중심결과에 대한 불분명한 효과	56

예시집 사용 관련 안내사항

AI 의료기술이 제공하는 가치를 종합적으로 평가하기 위해서는 다면적 가치평가가 필요합니다. 이 예시집은 AI가 제공하는 가치를 아래의 5가지 측면에서 평가하는 예시를 제공합니다.

가치 요소	수혜자	예시 및 설명
의학적 효과: 진단정확도*	환자	- 질병 진단/예측 성능: 민감도(sensitivity), 특이도(specificity), 수신자 조작 특성(receiver operating characteristic [ROC]) 곡선 면적 등
의학적 효과: 진료/환자 결과	환자	- 치료 효과/성공률, 생존율, 부작용 등
거시적 의료비용	보험자	- 특정 질환에 대한 국가/사회적 총 의료비용 - 의료서비스 사용 빈도(예: 영상 검사 횟수)
의료기관/의료인 효율	의료기관/의료인	- 의료기관의 업무 효율, 업무에 소요되는 비용 - 의료인의 진료 피로도, 업무 강도, 효율
비(非)임상 환자중심결과	환자	- 환자 경험, 만족도, 편의성(예: 조영제 사용량 감소, 검사 소요시간 단축, 방사선 조사 감소), 삶의 질

* 진단정확도 향상이 있는 경우, 아래의 방법을 이용하여 진료/환자 결과에 있어 net benefit을 초래할 수 있는 정도의 현저한/의미 있는 향상인지 그렇지 않은 약간의 개선에 불과한지에 대해서 좀 더 자세히 평가할 수 있음.

Net benefit

$$= \Delta\text{sensitivity} \times \text{prevalence} + \Delta\text{specificity} \times (1 - \text{prevalence}) \times \text{FP-to-TP outcome ratio}$$

FP: false positive, TP: true positive

다음의 웹 주소에서 net benefit 지표를 자동 계산하여 그래프를 생성할 수 있음

<https://aim-aicro.com/software/performanceoutcome>

참고문헌: Park SH, Sul AR, Han K, Sung YS. How to Determine If One Diagnostic Method, Such as an Artificial Intelligence Model, is Superior to Another: Beyond Performance Metrics. Korean J Radiol. 2023;24(7):601-605.

- 제공되는 예시는 건강보험 적용여부 평가를 받았던 실제 예들이 아니며, 단지 참조를 위한 예시로 동료평가를 거쳐 학술지에 게재된 논문들을 토대로 작성되었습니다. 따라서, 일부 AI 기술은 연구 단계로 아직 허가를 받지 않은 AI 의료기기가 포함되어 있습니다.
- 또한, 대부분의 예시는 한 편의 논문 결과를 토대로 작성되었습니다([예시 8]은 두 편의 논문 결과를 포함). 따라서, 각 예시는 각 예시에 포함된 AI에 대해 제시된 모든 연구결과를 전부 분석/평가한 것은 아니며 위의 5가지 가치 요소 중 일부만을 포함하고 있습니다.
- 하나의 AI 의료기술에 대하여 여러 연구결과가 있으면, 이들을 모두 평가하여 AI가 제공하는 다면적 가치를 아래와 같은 레이더 차트 형태로 종합할 수 있습니다. 대부분의 예시가 한 편의 논문 결과를 토대로 작성되었기 때문에, 제공되는 예시에는 이러한 종합 정리 및 차트는 생략되었습니다.



예시 1

현저한/의미 있는 진단정확도 향상, 진료/환자 결과 향상

예시 POINT

AI를 사용할 경우 AI를 사용하지 않는 진료에 비해 흉부 방사선사진의 판독에 있어 진단정확도가 향상되며, 이 결과를 이용해 진료/환자 결과에 있어 net benefit이 있는지 추정해 보았을 때 net benefit을 가져올 수 있는 현저한/의미 있는 진단정확도의 향상으로 판단됨. 또한, 연구를 통해 AI를 사용할 경우 AI를 사용하지 않는 통상의 진료에 비해 진료/환자 결과가 개선됨을 실제 데이터로 보여주었음.

가치 요소

- 의학적 효과: 진단정확도 향상
- 의학적 효과: 진료/환자 결과 향상

I. AI 모델

1. AI 모델 요약

1) 의료기기 인허가 자료 link	Lunit INSIGHT CXR version 2.0 [품목명] 2등급의료영상검출·진단보조소프트웨어/ [품목허가번호] 제 허 20-896호 / [최초허가일] 2020-10-19 / [유효기간] 2020-10-19 ~ 2025-10-18 [발급기관] 식품의약품안전처
2) 홈페이지 link	https://www.lunit.io/en/products/cxr
3) AI 모델의 기능/목적	흉부 방사선사진에서 주요 흉부 이상(폐결절, 폐렴, 기흉 등 - 제품에는 10가지로 되어있음[atelectasis, calcification, cardiomegaly, consolidation, fibrosis, mediastinal widening, nodule, pleural effusion, pneumoperitoneum, pneumothorax])을 감지하며, AI의 결과를 참고하여 의료인이 흉부 방사선사진을 판독함.
4) 사용자를 위한 AI 출력 결과 형태	주요 흉부 이상(폐결절, 폐렴, 기흉 등)의 가능성을 백분율 척도(0%~100%)로 평가하여 제시하며, 병변의 위치를 흉부 방사선사진 위에 heatmap으로 오버랩하여 표시함.
5) AI가 사용자에게 예측확률 또는 유사한 연속형 점수를 제시하는가?	<input type="checkbox"/> 제시하지 않음 <input type="checkbox"/> 제시하며, 기준 값 선택 및 점수 해석에 대한 명확한 지침을 제공함 <input checked="" type="checkbox"/> 제시하며, 기준 값 선택 및 점수 해석에 대한 명확한 지침을 제공하지 않음

2. 적용하려는 임상상황 및 대상환자

성인, 건강검진 대상 인구 (폐암 스크리닝)

3. 학습 데이터

1) 학습 데이터 기원	Lunit INSIGHT CXR version 2.0, 상업용 AI로 학습 데이터 set은 확인할 수 없음
2) 적용하려는 임상상황 및 대상환자와 부합하는가?	<input type="checkbox"/> 잘 부합함 <input checked="" type="checkbox"/> 불분명함 <input type="checkbox"/> 부합하지 않음
3) 국가	<input type="checkbox"/> 한국데이터 <input type="checkbox"/> 외국데이터 <input type="checkbox"/> 한국 및 외국데이터: 한국 (%), 외국 (%), <input checked="" type="checkbox"/> 비율 불분명

II. 임상 시험/검증 결과: 기존 진료와 AI를 활용한 진료 간 비교 결과

1. 의학적 효과 측면: 진단정확도 향상 여부

1) 자료 서지정보	Nam JG, Hwang EJ, Kim J, Park N, Lee EH, Kim HJ, et al. AI Improves Nodule Detection on Chest Radiographs in a Health Screening Population: A Randomized Controlled Trial. Radiology. 2023;307(2):e221894.
2) 시험/검증 데이터/환자 기원	2020년 6월 15일부터 2021년 12월 31일까지 건강검진 목적으로 3차 병원(서울대병원)을 방문하고 흉부 방사선촬영을 받은 19세 이상의 수검자 총 10,476명
3) 적용하려는 임상상황 및 대상환자와 부합하는가?	<input checked="" type="checkbox"/> 잘 부합함 <input type="checkbox"/> 불분명함 <input type="checkbox"/> 부합하지 않음(맞지 않는 환자군, 심한 selection bias)
4) 데이터/환자 수집 방법	<input type="checkbox"/> 후향적 <input checked="" type="checkbox"/> 전향적 <input type="checkbox"/> 전향적 및 후향적
5) 국가	<input checked="" type="checkbox"/> 한국데이터 <input type="checkbox"/> 외국데이터 <input type="checkbox"/> 한국 및 외국데이터: 한국 (%), 외국 (%), <input type="checkbox"/> 비율 불분명
6) 결과 요약	<p>핵심결과: 본 연구는 Randomized controlled trial (RCT)로 흉부 방사선사진을 촬영한 총 10,476명의 환자를 AI 보조 그룹(n = 5,238)과 AI 비보조 그룹(n = 5,238)으로 나누어 폐결절을 보고하였으며, 3개월 이내에 촬영한 후속 흉부 CT를 바탕으로 조치가 필요한(actionable) 폐결절(lung-RAD category 4 nodule: solid nodules ≥ 8 mm or subsolid nodules with a solid portion ≥ 6 mm)의 검출률을 보고자 한 연구임.</p> <p>연구의 일차 의료결과는 검출률이지만, 건강검진으로 시행한 검사이기 때문에 흉부 방사선사진의 결과와 관계없이 전체 코호트의 47% (n = 4,886)의 참가자에게 흉부 CT가 수행되었고 이에 대한 정확도 연구도 가능했음. 따라서, 흉부 CT를 시행한 subgroup을 대상으로 AI 그룹과 AI 비사용 그룹의 진단성능을 비교하는 연구가 추가로 수행되었으며, 두 군 사이의 sensitivity, positive predictive value, negative predictive value를 비교하였음.</p>

흉부 CT를 받은 참가자(n = 4,886)에서 조치가 필요한 폐결절의 진단에 있어 AI 그룹은 AI 비사용 그룹보다 높은 진단정확도를 보였음.

- sensitivity: 56.4% [31/55명] vs. 23.2% [13/56명], P < 0.001
- positive predictive value: 35.6% [31/87명] vs. 18.8% [13/69명], P = 0.02
- negative predictive value: 99.0% [2,314/2,338명] vs. 98.2% [2,349/2,392명], P = 0.03

Net benefit

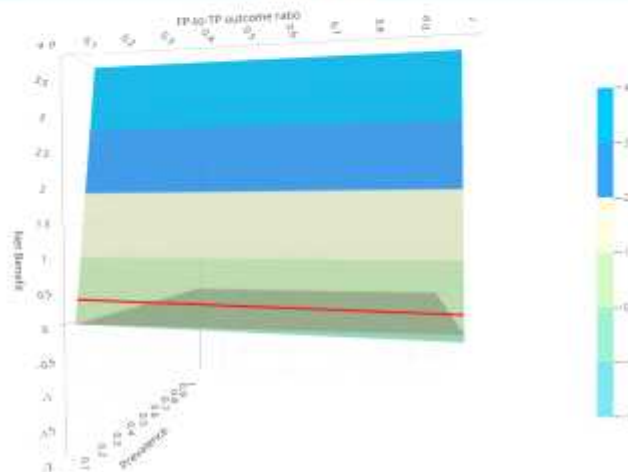
(<https://aim-aicro.com/software/performanceoutcome>) 분석:

- Δ Sensitivity: (33.2%)
- Δ Specificity: (-0.1%)
- Disease prevalence: 1.1% (10,476명의 참가자 중 111명, 폐결절 유병률)
- Net benefit 분석결과: 위에 계산한 값으로 Δ Sensitivity, Δ Specificity, Disease prevalence를 넣고, FP-to-TP outcome ratio를 0~1로 설정했을 때, FP-to-TP outcome ratio와 관계없이 net benefit이 발생하는 것으로 나타남.

Converting Performance Difference to Outcome Difference

Δ Sensitivity	33.2 %
Δ Specificity	-0.1 %
Disease prevalence	1.1 % to 1.1 %
FP-to-TP outcome ratio	0 to 1

Generate Graph



AI 효과 판정

- 도움됨
- 불분명
- 도움 안됨

2. 의학적 효과 측면: 진료/환자 결과 향상 여부

1) 자료 서지정보	Nam JG, Hwang EJ, Kim J, Park N, Lee EH, Kim HJ, et al. AI Improves Nodule Detection on Chest Radiographs in a Health Screening Population: A Randomized Controlled Trial. Radiology. 2023;307(2):e221894.	
2) 시험/검증 데이터/환자 기원	2020년 6월 15일부터 2021년 12월 31일까지 건강검진 목적으로 3차 병원(서울대병원)을 방문하고 흉부 방사선촬영을 받은 19세 이상의 환자 10,476명	
3) 적용하려는 임상상황 및 대상환자와 부합하는가?	<input checked="" type="checkbox"/> 잘 부합함 <input type="checkbox"/> 불분명함 <input type="checkbox"/> 부합하지 않음(맞지 않는 환자군, 심한 selection bias)	
4) 데이터/환자 수집 방법	<input type="checkbox"/> 후향적 <input checked="" type="checkbox"/> 전향적 <input type="checkbox"/> 전향적 및 후향적	
5) 국가	<input checked="" type="checkbox"/> 한국데이터 <input type="checkbox"/> 외국데이터 <input type="checkbox"/> 한국 및 외국데이터: 한국 (%), 외국 (%), <input type="checkbox"/> 비율 불분명	
6) 결과 요약	<p>핵심결과: RCT로 흉부 방사선사진을 촬영한 10,476명의 환자를 AI 보조 그룹(n = 5,238)과 AI 비사용 그룹(n = 5,238)으로 나누어 폐결절을 보고하였음. 3개월 이내에 촬영한 후속 흉부 CT를 바탕으로 조치가 필요한 폐결절(lung-RAD category 4 nodule: solid nodules ≥8 mm or subsolid nodules with a solid portion ≥6 mm)의 검출률(number of true-positive radiographs/total number of radiographs)을 비교함. 또한 두 그룹 사이의 false-referral rate, positive-report rate, 악성 폐결절 검출률, 폐암 검출률을 비교함.</p> <ul style="list-style-type: none"> - 조치가 필요한 폐결절 검출률: AI 그룹이 AI 비사용 그룹보다 높았음(0.59% [31/5,238명] vs. 0.25% [13/5,238명]; odds ratio, 2.4; 95% CI 1.3~4.7; P = 0.008). - false-referral rate: AI 그룹과 AI 비사용 그룹 간에 통계적으로 유의한 차이가 없었음(45.9% [56/122명] vs. 56.0% [56/100명]; P = 0.14). - positive-report rate: AI 그룹과 AI 비사용 그룹 간에 통계적으로 유의한 차이가 없었음(2.3% [122/5,238명] vs. 1.9% [100/5,238명]; P = 0.14). - 조직학적으로 확진된 악성 폐결절 검출률: AI 그룹에서 높았음(AI 그룹, 0.15% [8/5,238명] vs. AI 비사용 그룹, 0.0% [0/5,238명]; P = 0.008). - 폐암 검출률: AI 그룹에서 높았음(AI 그룹, 0.11% [6/5,238명] vs. AI 비사용 그룹, 0% [0/5,238명]; P = 0.03). 	<p>AI 효과 판정</p> <input checked="" type="checkbox"/> 도움됨 <input type="checkbox"/> 불분명 <input type="checkbox"/> 도움 안됨

3. 거시적 의료비용 감소여부

해당 결과 없음

4. 의료기관/의료인 효율 향상 여부

해당 결과 없음

5. 환자중심결과 향상 측면

해당 결과 없음

예시 2

의료기관/의료인 효율 향상

예시 POINT

AI를 사용하였을 때 AI를 사용하지 않는 통상의 진료에 비해 손 방사선사진을 이용한 골연령 판독에 있어 의료인의 효율을 높일 수 있음을 보여준 예

가치 요소

- 의료기관/의료인 효율 향상

I. AI 모델

1. AI 모델 요약

1) 의료기기 인허가 자료 link	CE 인증 (2019), ISO13485:2016certification, KFDA approval, EU registration
2) 홈페이지 link	http://medi.ai.io/
3) AI 모델의 기능/목적	손 방사선사진에서 7가지 부위를 발견, 골 성숙도를 분류하여, 두 가지의 영상의학적 골연령을 제시하여(hybrid Greulich-Pyle [GP] 방법과 modified Tanner-Whitehouse [TW] 방법), 의료인이 골연령 측정을 위한 참고자료로 사용함.
4) 사용자를 위한 AI 출력 결과 형태	7가지 부위(radius, ulna, distal phalange, middle phalange, proximal phalange, metacarpal of the third digit, metacarpal of the first digit)의 골 성숙도를 1.5세~18세까지 6개월 간격으로 제시함. 골 성숙도 분석 위치를 손 방사선사진 위에 heatmap으로 오버랩하여 표시함.
5) AI가 사용자에게 예측확률 또는 유사한 연속형 점수를 제시하는가?	<input checked="" type="checkbox"/> 제시하지 않음 <input type="checkbox"/> 제시하며, 기준 값 선택 및 점수 해석에 대한 명확한 지침을 제공함 <input type="checkbox"/> 제시하며, 기준 값 선택 및 점수 해석에 대한 명확한 지침을 제공하지 않음

2. 적용하려는 임상상황 및 대상환자

18세 이하의 소아 환자, 특히 내분비 질환이 있는 소아 환자의 경과나 치료 평가 및 성인 키 예측 등에 활용될 수 있음.

3. 학습 데이터

1) 학습 데이터 기원	두 개의 공공 데이터셋으로부터 기원함. - Society of North America (RSNA) 2017 challenge: 미국 Stanford University와 Colorado University에서 얻은 14,236개의 손 방사선사진 - Digital Hand Atlas: 미국 University of Southern California에서 얻은 1,375개의 손 방사선사진
2) 적용하려는 임상상황 및 대상환자와 부합하는가?	<input checked="" type="checkbox"/> 잘 부합함 <input type="checkbox"/> 불분명함 <input type="checkbox"/> 부합하지 않음
3) 국가	<input type="checkbox"/> 한국데이터 <input checked="" type="checkbox"/> 외국데이터 <input type="checkbox"/> 한국 및 외국데이터: 한국(%), 외국 (%), <input type="checkbox"/> 비율 불분명

II. 임상 시험/검증 결과: 기존 진료와 AI를 활용한 진료 간 비교 결과

1. 의학적 효과 측면: 진단정확도 향상 여부

해당 결과 없음

2. 의학적 효과 측면: 진료/환자 결과 향상 여부

해당 결과 없음

3. 거시적 의료비용 감소 여부

해당 결과 없음

4. 의료기관/의료인 효율 향상 여부

1) 자료 서지정보	Lee KC, Lee KH, Kang CH, Ahn KS, Chung LY, Lee JJ, et al. Clinical Validation of a Deep Learning-Based Hybrid (Greulich-Pyle and Modified Tanner-Whitehouse) Method for Bone Age Assessment. Korean J Radiol. 2021;22(12):2017-2025.	
2) 시험/검증 데이터/환자 기원	2017년 6월부터 2018년 8월까지 고려대 안암병원에 방문한 2세~18세 소아 환자 102명의 손 방사선사진 영상 1,646개	
3) 적용하려는 임상상황 및 대상환자와 부합하는가?	<input checked="" type="checkbox"/> 잘 부합함 <input type="checkbox"/> 불분명함 <input type="checkbox"/> 부합하지 않음(맞지 않는 환자군, 심한 selection bias)	
4) 데이터/환자 수집 방법	<input checked="" type="checkbox"/> 후향적 <input type="checkbox"/> 전향적 <input type="checkbox"/> 전향적 및 후향적	
5) 국가	<input checked="" type="checkbox"/> 한국데이터 <input type="checkbox"/> 외국데이터 <input type="checkbox"/> 한국 및 외국데이터: 한국(%), 외국 (%), <input type="checkbox"/> 비율 불분명	
6) 결과 요약	<p>핵심결과: 두 명의 영상의학자가 두 번의 separate session에서 AI의 보조를 받거나 받지 않고 골연령을 판독하고 reading time이 기록되었고, AI의 보조를 받거나 받지 않은 경우에 reading time의 차이를 구함. AI를 이용하여 골연령을 평가 시 두 판독자의 reading time이 평균적으로 54.29초에서 35.37초로 18.6초 (95% CI, 12.9초~24.3초) 감소함(P < 0.001).</p>	<p>AI 효과 판정 <input checked="" type="checkbox"/> 도움됨 <input type="checkbox"/> 불분명 <input type="checkbox"/> 도움 안됨</p>

5. 환자중심결과 향상 측면

해당 결과 없음

예시 3

의료기관/의료인 효율 향상, 환자중심결과 향상

예시 POINT

AI를 사용하였을 때 AI를 사용하지 않는 통상의 진료에 비해 무릎 부위 MRI 촬영시간을 단축시켜 의료업무의 효율 및 환자의 편의성 향상을 보여주는 예

가치 요소

- 의료기관/의료인 효율 향상
- 환자중심결과 향상

I. AI 모델

1. AI 모델 요약

1) 의료기기 인허가 자료 link	의료기기 허가 받지 않음
2) 홈페이지 link	The Fundamental AI Research (FAIR) team at META에 소속되어 있다고 밝힘. https://ai.meta.com/research/ Code 공개: https://github.com/facebookresearch/fastMRI/tree/main/fastmri_examples/RadiologyJohnson2022
3) AI 모델의 기능/목적	3-T MRI로 knee MRI를 촬영 시 AI를 이용하여 영상의 질을 유지하면서 촬영 시간을 단축
4) 사용자를 위한 AI 출력 결과 형태	AI로 빠르게 촬영하면서 화질이 개선된 영상
5) AI가 사용자에게 예측확률 또는 유사한 연속형 점수를 제시하는가?	<input checked="" type="checkbox"/> 제시하지 않음 <input type="checkbox"/> 제시하며, 기준 값 선택 및 점수 해석에 대한 명확한 지침을 제공함 <input type="checkbox"/> 제시하며, 기준 값 선택 및 점수 해석에 대한 명확한 지침을 제공하지 않음

2. 적용하려는 임상상황 및 대상환자

임상적으로 진단 목적의 knee MRI가 필요한 환자

3. 학습 데이터

1) 학습 데이터 기원	NYU Langone Health에서 2019년 2월~3월, commercial 3-T MRI systems (MAGNETOM Skyra and Biograph mMR, Siemens Healthineers)을 이용하여 촬영된 298개의 knee MRI
2) 적용하려는 임상상황 및 대상환자와 부합하는가?	<input type="checkbox"/> 잘 부합함 <input checked="" type="checkbox"/> 불분명함 <input type="checkbox"/> 부합하지 않음
3) 국가	<input type="checkbox"/> 한국데이터 <input checked="" type="checkbox"/> 외국데이터 <input type="checkbox"/> 한국 및 외국데이터: 한국 (%), 외국 (%), <input type="checkbox"/> 비율 불분명

II. 임상 시험/검증 결과: 기존 진료와 AI를 활용한 진료간 비교 결과

1. 의학적 효과 측면: 진단정확도 향상 여부

해당 결과 없음

2. 의학적 효과 측면: 진료/환자 결과 향상 여부

해당 결과 없음

3. 거시적 의료비용 감소 여부

해당 결과 없음

4. 의료기관/의료인 효율 향상 여부

1) 자료 서지정보	Johnson PM, Lin DJ, Zbontar J, Zitnick CL, Sriram A, Muckley M, et al. Deep Learning Reconstruction Enables Prospectively Accelerated Clinical Knee MRI. Radiology. 2023;307(2):e220425.
2) 시험/검증 데이터/환자 기원	NYU Langone Health, 2020년 1월 ~ 2021년 2월, 진단 목적의 knee MRI를 얻으려는 18세 이상의 성인 참가자 중 금속 인공물이 있거나, MRI의 금기증에 해당하거나, knee MRI coil에 적합하지 않은 무릎 크기를 가진 환자를 제외한 총 170명의 참가자들에서 3-T MRI system (MAGNETOM Skyra, Siemens Healthineers)을 이용하여 얻어진 knee MRI
3) 적용하려는 임상상황 및 대상환자와 부합하는가?	<input checked="" type="checkbox"/> 잘 부합함 <input type="checkbox"/> 불분명함 <input type="checkbox"/> 부합하지 않음(맞지 않는 환자군, 심한 selection bias)
4) 데이터/환자 수집 방법	<input type="checkbox"/> 후향적 <input checked="" type="checkbox"/> 전향적 <input type="checkbox"/> 전향적 및 후향적
5) 국가	<input type="checkbox"/> 한국데이터 <input checked="" type="checkbox"/> 외국데이터 <input type="checkbox"/> 한국 및 외국데이터: 한국 (%), 외국 (%), <input type="checkbox"/> 비율 불분명
6) 결과 요약	핵심결과: AI를 이용한 deep learning (DL) protocol은 평균 촬영시 간이평균 5분33초 ±16초로, 기존 촬영시간 9분56초 ±19초에서 단축되었음. Scan time의 감축은 MRI 검사시간을 단축시켜서 병원 진료 업무의 효율을 높이고 검사 환자 수를 증가 시킬 수 있으며, 환자가 MRI 기기 안에 누워 있어야 하는 시간을 줄여 환자의 불편을 줄일 수 있음.

5. 환자중심결과 향상 측면

1) 자료 서지정보	Johnson PM, Lin DJ, Zbontar J, Zitnick CL, Sriram A, Muckley M, et al. Deep Learning Reconstruction Enables Prospectively Accelerated Clinical Knee MRI. Radiology. 2023;307(2):e220425.	
2) 시험/검증 데이터/환자 기원	NYU Langone Health, 2020년 1월 ~ 2021년 2월, 진단 목적의 knee MRI를 얻으려는 18세 이상의 성인 참가자 중 금속 인공물이 있거나, MRI의 금기증에 해당하거나, knee MRI coil에 적합하지 않은 무릎 크기를 가진 환자를 제외한 총 170명의 참가자들에서 3-T MRI system (MAGNETOM Skyra, Siemens Healthineers)을 이용하여 얻어진 knee MRI	
3) 적용하려는 임상상황 및 대상환자와 부합하는가?	<input checked="" type="checkbox"/> 잘 부합함 <input type="checkbox"/> 불분명함 <input type="checkbox"/> 부합하지 않음(맞지 않는 환자군, 심한 selection bias)	
4) 데이터/환자 수집 방법	<input type="checkbox"/> 후향적 <input checked="" type="checkbox"/> 전향적 <input type="checkbox"/> 전향적 및 후향적	
5) 국가	<input type="checkbox"/> 한국데이터 <input checked="" type="checkbox"/> 외국데이터 <input type="checkbox"/> 한국 및 외국데이터: 한국 (%), 외국 (%), <input type="checkbox"/> 비율 불분명	
6) 결과 요약	<p>핵심결과: AI를 이용한 deep learning (DL) protocol은 평균 촬영시간이평균 5분33초±16초로, 기존 촬영시간 9분56초±19초에서 단축되었음. Scan time의 감축은 MRI 검사시간을 단축시켜서 병원 진료 업무의 효율을 높이고 검사 환자 수를 증가시킬 수 있으며, 환자가 MRI 기기 안에 누워 있어야 하는 시간을 줄여 환자의 불편을 줄일 수 있음.</p>	<p>AI 효과 판정 <input checked="" type="checkbox"/> 도움됨 <input type="checkbox"/> 불분명 <input type="checkbox"/> 도움 안됨</p>

예시 4

거시적 의료비용 감소

예시 POINT

후향적 simulation 연구를 통해, AI를 사용하였을 때 AI를 사용하지 않는 통상의 진료보다 폐암 검진에 있어 불필요한 검진 대상자를 감소시켜 거시적 의료비 절감의 가능성을 보여준 예. 다만, 의료비 절감의 효과를 실제 직접적으로 평가하지는 않았음.

가치 요소

- 거시적 의료비용 감소

I. AI 모델

1. AI 모델 요약

1) 의료기기 인허가 자료 link	의료기기 허가 받지 않음
2) 홈페이지 link	Code 공개: https://github.com/vineet1992/CXR-LC
3) AI 모델의 기능/목적	임상 정보(나이, 성별, 흡연 유무)와 흉부 방사선사진을 이용하여 폐암 발생 고 위험군, 즉 저선량 CT 폐암 선별검사의 적절한 후보자를 선정하고자 함.
4) 사용자를 위한 AI 출력 결과 형태	12년 이내에 폐암이 발생할 위험에 대한 확률(0%~100%)을 계산하여, 4개의 위험 범주로 분류함: (a) low risk (< 2%), (b) indeterminate risk (2% ~ < 3.297%), (c) high risk (3.297% ~ < 8%), (d) very high risk (≥ 8%)
5) AI가 사용자에게 예측확률 또는 유사한 연속형 점수를 제시하는가?	<input type="checkbox"/> 제시하지 않음 <input checked="" type="checkbox"/> 제시하며, 기준 값 선택 및 점수 해석에 대한 명확한 지침을 제공함 <input type="checkbox"/> 제시하며, 기준 값 선택 및 점수 해석에 대한 명확한 지침을 제공하지 않음

2. 적용하려는 임상상황 및 대상환자

건강검진 대상인구(폐암 스크리닝), 특히 임상적으로 폐암 위험인자(고령, 흡연력)가 있는 성인

3. 학습 데이터

1) 학습 데이터 기원	미국 10개 검진병원에서 1993년부터 2001년까지 The Prostate, Lung, Colorectal, and Ovarian (PLCO) Randomized Trial에 등록된 암 검진 참가자 총 41,856명의 흉부 x선
2) 적용하려는 임상상황 및 대상환자와 부합하는가?	<input type="checkbox"/> 잘 부합함 <input checked="" type="checkbox"/> 불분명함 <input type="checkbox"/> 부합하지 않음
3) 국가	<input type="checkbox"/> 한국데이터 <input checked="" type="checkbox"/> 외국데이터 <input type="checkbox"/> 한국 및 외국데이터: 한국 (%), 외국 (%), <input type="checkbox"/> 비율 불분명

II. 임상 시험/검증 결과: 기존 진료와 AI를 활용한 진료 간 비교 결과

1. 의학적 효과 측면: 진단정확도 향상 여부

해당 결과 없음

2. 의학적 효과 측면: 진료/환자 결과 향상 여부

해당 결과 없음

3. 거시적 의료비용 감소 여부

1) 자료 서지정보	Lee JH, Lee D, Lu MT, Raghu VK, Park CM, Goo JM, et al. Deep Learning to Optimize Candidate Selection for Lung Cancer CT Screening: Advancing the 2021 USPSTF Recommendations. Radiology. 2022;305(1): 209-218.
2) 시험/검증 데이터/환자 기원	서울대학교 강남센터(검진 병원), 2004년 1월 ~ 2018년 6월에 얻어진, 흡연력 정보가 있는 50세~80세 성인의 건강검진 흉부 방사선촬영 19,488건
3) 적용하려는 임상상황 및 대상환자와 부합하는가?	<input checked="" type="checkbox"/> 잘 부합함 <input type="checkbox"/> 불분명함 <input type="checkbox"/> 부합하지 않음(맞지 않는 환자군, 심한 selection bias)
4) 데이터/환자 수집 방법	<input checked="" type="checkbox"/> 후향적 <input type="checkbox"/> 전향적 <input type="checkbox"/> 전향적 및 후향적
5) 국가	<input checked="" type="checkbox"/> 한국데이터 <input type="checkbox"/> 외국데이터 <input type="checkbox"/> 한국 및 외국데이터: 한국 (%), 외국 (%), <input type="checkbox"/> 비율 불분명

<p>6) 결과 요약</p>	<p>핵심결과: 폐암 CT 선별검사의 후보자를 선택하는 데에 있어서 흉부 x선 AI 위험모델을 추가로 이용하는 것이 검사의 선별검사 대상자 수를 줄이는데 도움이 되는지를 평가함.</p> <p>2021년 U.S. Preventive Services Task Force (USPSTF) 권고에 부합하는 경우(50세~80세, 20갑년 이상의 흡연력이 있는 현재 흡연자 또는 15년 이내에 금연한 자) vs. USPSTF에 부합하는 경우 중 AI를 이용한 검진 흉부 x선에서 폐암 발생의 위험도가 low~intermediate인 경우(위험 < 3.297%)를 제외했을 때를 비교함.</p> <p>AI를 추가로 이용하는 경우에 CT 선별검사의 후보자 수는 통계적으로 유의미하게 감소함.</p> <ul style="list-style-type: none"> - CT 선별검사 후보자 비율: 45.1% (7,835/17,390) vs. 35.8% (6,233/17,390), P < 0.001 - AI를 추가로 이용하면 USPSTF 기준만 적용 시보다 20.4% (1,602/7,835)의 선별검사 대상 숫자를 줄일 수 있음. <p>두 그룹 간 폐암 포함율(0.3% [53/17,390] vs 0.3% [56/17,390], P = 0.85), positive predictive value (0.9% vs 0.7%, P = 0.42)는 통계적으로 유의미한 차이를 보이지 않았음.</p>	<p>AI 효과 판정</p> <p><input checked="" type="checkbox"/> 도움됨</p> <p><input type="checkbox"/> 불분명</p> <p><input type="checkbox"/> 도움 안됨</p>
-----------------	--	---

4. 의료기관/의료인 효율 향상 여부

해당 결과 없음

5. 환자중심결과 향상 측면

해당 결과 없음

예시 5

진료/환자 결과 향상

예시 POINT

전향적 RCT 연구를 통해, AI를 사용할 경우 그렇지 않을 경우에 비해 전신 마취 하 수술 시 환자의 저혈압 발생을 감소시킬 수 있음을 보여준 예

가치 요소

- 의학적 효과: 진료/환자 결과 향상

I. AI 모델

1. AI 모델 요약

1) 의료기기 인허가 자료 link	510(k), K203224 Trade/Device Name: Acumen™ Hypotension Prediction Index Regulation Number: 21 CFR 870.2210 Regulation Name: Adjunctive predictive cardiovascular indicator Regulatory Class: Class II Product Code: QAQ https://assets-us-01.kc-usercontent.com/6239a81e-8f0f-0040-a1df-b4932a10f6ae/21b36895-0bca-4c0c-be21-1dc7e1d65f2d/K203224.pdf
2) 홈페이지 link	https://www.edwards.com/healthcare-professionals/products-services/predictive-monitoring/hpi
3) AI 모델의 기능/목적	전신 마취를 한 상태에서 수술을 받는 환자에게 수술 중 발생하는 저혈압(평균 동맥압 < 65mm Hg)은 합병증 및 사망률 증가와 관련 있는 것으로 알려져 있음. 해당 AI 모델은 일반적으로 전신 마취 시에 마취과 의사가 모니터링하는 동맥 파형에 대한 데이터를 머신러닝으로 학습하여, 실제 저혈압이 발생하기 전에 이를 미리 예상할 수 있는 경고를 제공하는 알고리즘임.
4) 사용자를 위한 AI 출력 결과 형태	본 AI 모델은 환자의 동맥 파형을 연속적으로 측정하면서, 이를 바탕으로 환자에게 15분 내로 저혈압이 발생할 확률을 20초마다 제공함. 저혈압이 발생하기까지의 시간 예측은 제공하지 않으며, 저혈압 발생 확률이 85%보다 높은 경우 알람 소리와 깜박이는 빛이 발생함. 이와 함께, 예상되는 저혈압의 원인에 대한 정보를 제공함(심박수, 심박출량, 시간에 따른 압력 변화, 일회박출량, 일회박출량 변화, 혈관 저항).
5) AI가 사용자에게 예측확률 또는 유사한 연속형 점수를 제시하는가?	<input type="checkbox"/> 제시하지 않음 <input checked="" type="checkbox"/> 제시하며, 기준 값 선택 및 점수 해석에 대한 명확한 지침을 제공함 <input type="checkbox"/> 제시하며, 기준 값 선택 및 점수 해석에 대한 명확한 지침을 제공하지 않음

2. 적용하려는 임상상황 및 대상환자

- 전신 마취 하에서 응급이 아닌 선택적 수술을 받는 성인(만 18세 이상) 환자
- 심장 관련 수술 환자는 제외
 - 평소 저혈압이 있는 환자는 제외

3. 학습 데이터

1) 학습 데이터 기원	미국 보스턴의 MIMIC database에 포함된 중환자실 환자(326명) 및 미국 Edwards 데이터베이스에 포함된 중환자실 환자(688명) 및 수술 환자(670명)
2) 적용하려는 임상상황 및 대상환자와 부합하는가?	<input type="checkbox"/> 잘 부합함 <input type="checkbox"/> 불분명함 <input checked="" type="checkbox"/> 부합하지 않음
3) 국가	<input type="checkbox"/> 한국데이터 <input checked="" type="checkbox"/> 외국데이터 <input type="checkbox"/> 한국 및 외국데이터: 한국 (%), 외국 (%), <input type="checkbox"/> 비율 불분명

II. 임상 시험/검증 결과: 기존 진료와 AI를 활용한 진료 간 비교 결과

1. 의학적 효과 측면: 진단정확도 향상 여부

해당 결과 없음

2. 의학적 효과 측면: 진료/환자 결과 향상 여부

1) 자료 서지정보	Wijnberge M, Geerts BF, Hol L, Lemmers N, Mulder MP, Berge P, et al. Effect of a Machine Learning-Derived Early Warning System for Intraoperative Hypotension vs Standard Care on Depth and Duration of Intraoperative Hypotension During Elective Noncardiac Surgery: The HYPE Randomized Clinical Trial. JAMA. 2020;323(11):1052-1060.	
2) 시험/검증 데이터/환자 기원	네덜란드 암스테르담의 3차 병원 한 곳에서 2018년 5월 ~ 2019년 3월에 전신 마취 하에 응급이 아닌 선택적 수술을 받는 성인 환자 68명. 기존 수술 환자 모니터링에 AI 모델을 같이 적용한 그룹(중재군)과 기존 모니터링만을 사용한 그룹(대조군)에 각각 34 명씩 배정되었음.	
3) 적용하려는 임상상황 및 대상환자와 부합 하는가?	<input checked="" type="checkbox"/> 잘 부합함 <input type="checkbox"/> 불분명함 <input type="checkbox"/> 부합하지 않음(맞지 않는 환자군, 심한 selection bias)	
4) 데이터/환자 수집 방법	<input type="checkbox"/> 후향적 <input checked="" type="checkbox"/> 전향적 <input type="checkbox"/> 전향적 및 후향적	
5) 국가	<input type="checkbox"/> 한국데이터 <input checked="" type="checkbox"/> 외국데이터 <input type="checkbox"/> 한국 및 외국데이터: 한국 (%), 외국 (%), <input type="checkbox"/> 비율 불분명	
6) 결과 요약	<p>핵심결과: Unblinded RCT로 parallel design</p> <ul style="list-style-type: none"> - 일차 결과 값은 수술 중 발생한 저혈압의 median time-weighted average이었음. 중재군 및 대조군에서 각각 0.10 mmHg vs. 0.44 mmHg (P = 0.001)로, AI 모델을 같이 사용할 경우에 수술 중 저혈압의 발생 정도가 줄어들었음을 알 수 있음. - 이차 결과 값으로 AI 모델을 사용하였을 때의 과잉 치료를 받을 가능성을 평가하기 위하여 두 그룹 간에 고혈압 (평균 동맥압 > 100 mm Hg)의 median time-weighted average와 저혈압 치료제의 누적 사용량 등을 비교하였으나, 유의미한 차이는 없었음. 따라서 해당 AI 모델을 추가로 사용하여 전신 마취 환자를 모니터링 할 경우에 수술 중 발생하는 저혈압의 빈도를 줄일 수 있으며, 유의미한 과잉 치료는 발생하지 않는 것으로 판단됨. 	<p>AI 효과 판정</p> <input checked="" type="checkbox"/> 도움됨 <input type="checkbox"/> 불분명 <input type="checkbox"/> 도움 안됨

3. 거시적 의료비용 감소 여부

해당 결과 없음

4. 의료기관/의료인 효율 향상 여부

해당 결과 없음

5. 환자중심결과 향상 측면

해당 결과 없음

예시 6

진료/환자 결과에 대한 불분명한 효과, 의료기관/의료인 효율 감소

예시 POINT

RCT 연구를 통해, AI를 사용할 경우 그렇지 않은 경우에 비해 1) 대장내시경을 이용한 대장암 검진의 결과(선종 검출률)를 향상시킬 가능성을 보여주었으나 통계적 유의성은 없었고, 2) 대장내시경검사에 소요되는 시간은 유의하게 증가되어 의료기관/의료인의 효율은 감소되는 것을 보여준 예

가치 요소

- 의학적 효과: 진료/환자 결과에 대한 불분명한 효과
- 의료기관/의료인 효율 감소

I. AI 모델

1. AI 모델 요약

1) 의료기기 인허가 자료 link	510(k), K211951 Trade/Device Name: GI Genius Regulation Number: 21 CFR 876.1520 Regulation Name: Gastrointestinal Lesion Software Detection System Regulatory Class: Class II Product Code: QNP https://www.accessdata.fda.gov/cdrh_docs/pdf21/K211951.pdf
2) 홈페이지 link	https://www.medtronic.com/covidien/en-gb/products/gastrointestinal-artificial-intelligence/gi-genius-intelligent-endoscopy.html#
3) AI 모델의 기능/목적	대장내시경을 시행할 때 내시경 processor로부터 digital image를 분석하여, 점막 병변(용종, 선종)이 의심될 경우 실시간으로 녹색 박스를 해당 병변 주변에 출력하여 주는 프로그램임. AI 모델의 출력 결과를 대장내시경을 시행하는 의사가 참고하여 용종 발견에 활용함.
4) 사용자를 위한 AI 출력 결과 형태	대장내시경의 digital image를 실시간으로 분석하여, 용종이 의심되는 경우 녹색 박스를 해당 병변 주변에 출력함.
5) AI가 사용자에게 예측확률 또는 유사한 연속형 점수를 제시하는가?	<input checked="" type="checkbox"/> 제시하지 않음 <input type="checkbox"/> 제시하며, 기준 값 선택 및 점수 해석에 대한 명확한 지침을 제공함 <input type="checkbox"/> 제시하며, 기준 값 선택 및 점수 해석에 대한 명확한 지침을 제공하지 않음

2. 적용하려는 임상상황 및 대상환자

무증상의 평균적인 위험을 가진 성인(만 18세 이상)을 대상으로 하는 검진 대장내시경검사

3. 학습 데이터

1) 학습 데이터 기원	2013년 12월부터 2016년 10월까지 유럽과 미국에서 검진 대장내시경을 시행한 50세~75세의 환자 중 적어도 한 개 이상의 용종이 있으며 기록된 내시경 영상에 기술적 문제가 없는 환자(총 840명; 병변 2,684개).
2) 적용하려는 임상상황 및 대상환자와 부합하는가?	<input checked="" type="checkbox"/> 잘 부합함 <input type="checkbox"/> 불분명함 <input type="checkbox"/> 부합하지 않음
3) 국가	<input type="checkbox"/> 한국데이터 <input checked="" type="checkbox"/> 외국데이터 <input type="checkbox"/> 한국 및 외국데이터: 한국 (%), 외국 (%), <input type="checkbox"/> 비율 불분명

II. 임상 시험/검증 결과: 기존 진료와 AI를 활용한 진료 간 비교 결과

1. 의학적 효과 측면: 진단정확도 향상 여부

해당 결과 없음

2. 의학적 효과 측면: 진료/환자 결과 향상 여부

1) 자료 서지정보	Karsenti D, Tharsis G, Perrot B, Cattan P, Percie du Sert A, Venezia F, et al. Effect of real-time computer-aided detection of colorectal adenoma in routine colonoscopy (COLO-GENIUS): a single-centre randomised controlled trial. <i>Lancet Gastroenterol Hepatol.</i> 2023;8(8):726-734.	
2) 시험/검증 데이터/환자 기원	2021년 5월 1일부터 2022년 5월 1일까지 프랑스의 내시경 센터 한 곳에서 검진 목적으로 대장내시경을 시행한 성인 환자. Unblinded RCT, parallel design으로 기존 표준 대장내시경을 시행 받은 그룹(대조군)과 AI를 적용한 그룹(실험군)에 각각 1,026명 및 1,013명씩 배정되었음.	
3) 적용하려는 임상상황 및 대상환자와 부합하는가?	<input checked="" type="checkbox"/> 잘 부합함 <input type="checkbox"/> 불분명함 <input type="checkbox"/> 부합하지 않음(맞지 않는 환자군, 심한 selection bias)	
4) 데이터/환자 수집 방법	<input type="checkbox"/> 후향적 <input checked="" type="checkbox"/> 전향적 <input type="checkbox"/> 전향적 및 후향적	
5) 국가	<input type="checkbox"/> 한국데이터 <input checked="" type="checkbox"/> 외국데이터 <input type="checkbox"/> 한국 및 외국데이터: 한국 (%), 외국 (%), <input type="checkbox"/> 비율 불분명	
6) 결과 요약	핵심결과: 일차 결과 값은 한 개 이상의 선종이 확인된 대장내시경검사의 비율 (adenoma detection rate [ADR])로 표준 대장내시경 군에서는 33.7% (341/1,012), AI 적용군에서는 37.5% (376/1,003) (P = 0.051)로 통계적 유의미함은 경계선에 있음.	AI 효과 판정 <input type="checkbox"/> 도움됨 <input checked="" type="checkbox"/> 불분명 <input type="checkbox"/> 도움 안됨

3. 거시적 의료비용 감소 여부

해당 결과 없음

4. 의료기관/의료인 효율 향상 여부

1) 자료 서지정보	Karsenti D, Tharsis G, Perrot B, Cattan P, Percie du Sert A, Venezia F, et al. Effect of real-time computer-aided detection of colorectal adenoma in routine colonoscopy (COLO-GENIUS): a single-centre randomised controlled trial. <i>Lancet Gastroenterol Hepatol.</i> 2023;8(8):726-734.	
2) 시험/검증 데이터/환자 기원	2021년 5월 1일부터 2022년 5월 1일까지 프랑스의 내시경 센터 한 곳에서 검진 목적으로 대장내시경을 시행한 성인 환자. 기존 표준 대장내시경을 시행 받은 그룹(대조군)과 AI를 적용한 그룹(실험군)에 각각 1,026명 및 1,013명씩 배정되었음.	
3) 적용하려는 임상상황 및 대상환자와 부합하는가?	<input checked="" type="checkbox"/> 잘 부합함 <input type="checkbox"/> 불분명함 <input type="checkbox"/> 부합하지 않음(맞지 않는 환자군, 심한 selection bias)	
4) 데이터/환자 수집 방법	<input type="checkbox"/> 후향적 <input checked="" type="checkbox"/> 전향적 <input type="checkbox"/> 전향적 및 후향적	
5) 국가	<input type="checkbox"/> 한국데이터 <input checked="" type="checkbox"/> 외국데이터 <input type="checkbox"/> 한국 및 외국데이터: 한국 (%), 외국 (%), <input type="checkbox"/> 비율 불분명	
6) 결과 요약	<p>핵심결과: 이차 결과 값으로 대장탐색시간(colonic exploration time)과 대장후퇴시간(colonic withdrawal time)을 계산하였음. 이는 검사 소요시간 지표로, 의료인의 효율성을 평가할 수 있는 지표임. 표준내시경을 시행 받은 군이 AI 적용군과 비교 시 대장탐색시간과 대장후퇴시간이 모두 유의하게 짧았음. - 대장탐색시간: 510초 (interquartile range [IQR] 392초~700초) vs. 570초 (IQR 420초~792초) (P < 0.0001) - 대장후퇴시간: 435초 (IQR 360초~583초) vs. 480초 (IQR 360초~648초) (P = 0.0013) 이를 통해 AI를 사용할 때 검사 당 소요되는 시간은 증가하여, 의료인 효율성은 떨어지는 것으로 판단됨.</p>	<p>AI 효과 판정</p> <input type="checkbox"/> 도움됨 <input type="checkbox"/> 불분명 <input checked="" type="checkbox"/> 도움 안됨

5. 환자중심결과 향상 측면

해당 결과 없음

예시 7

의료기관/의료인 효율 향상

예시 POINT

AI를 사용할 경우 그렇지 않은 경우에 비해 방사선치료를 위한 치료 계획(방사선 조사에 필요한 병변 부위 윤곽 그리기) 시간을 93% 단축시켜, 진료의 효율을 높일 수 있음을 보여준 예

가치 요소

- 의료기관/의료인 효율 향상

I. AI 모델

1. AI 모델 요약

1) 의료기기 인허가 자료 link	의료기기 인허가 이전(Microsoft의 Open source software로 만들어진 AI로, 상업화되지 않은 것으로 생각됨)
2) 홈페이지 link	https://www.microsoft.com/en-us/research/project/medical-image-analysis/ https://news.microsoft.com/en-gb/2023/06/27/ai-helping-shrink-waiting-times-nhs-cancer-patients/
3) AI 모델의 기능/목적	- 기능: 자동 장기 분할(autocontouring solution) - 목적: 전립선암, 두경부암의 방사선치료 계획을 위해 표적 장기와 주변 위험 장기(organs at risk [OAR])의 윤곽을 자동으로 분할함.
4) 사용자를 위한 AI 출력 결과 형태	관심 장기/병변을 자동 분할
5) AI가 사용자에게 예측확률 또는 유사한 연속형 점수를 제시하는가?	<input checked="" type="checkbox"/> 제시하지 않음 <input type="checkbox"/> 제시하며, 기준 값 선택 및 점수 해석에 대한 명확한 지침을 제공함 <input type="checkbox"/> 제시하며, 기준 값 선택 및 점수 해석에 대한 명확한 지침을 제공하지 않음

2. 적용하려는 임상상황 및 대상환자

알고 있는 두경부암 또는 전립선암에 대한 방사선치료를 시작하려는 환자

3. 학습 데이터

1) 학습 데이터 기원	- 골반 CT: 2012년~2019년까지 5개 병원에서 전립선암으로 방사선치료를 받은 남자 환자 CT 총 345개(유럽, 호주, 뉴질랜드, 북미 - 병원 이름은 자세히 언급되어 있지 않음) - 두경부 CT: 2개 병원에서 두경부암으로 방사선치료를 받은 환자(여성 24%, 남성 76%) CT 총 176개(유럽, 북미 - 전립선암 환자 CT 데이터를 얻어온 병원과 동일)
2) 적용하려는 임상상황 및 대상환자와 부합하는가?	<input checked="" type="checkbox"/> 잘 부합함 <input type="checkbox"/> 불분명함 <input type="checkbox"/> 부합하지 않음
3) 국가	<input type="checkbox"/> 한국데이터 <input checked="" type="checkbox"/> 외국데이터 <input type="checkbox"/> 한국 및 외국데이터: 한국 (%), 외국 (%), <input type="checkbox"/> 비율 불분명

II. 임상 시험/검증 결과: 기존 진료와 AI를 활용한 진료 간 비교 결과

1. 의학적 효과 측면: 진단정확도 향상 여부

해당 결과 없음

2. 의학적 효과 측면: 진료/환자 결과 향상 여부

해당 결과 없음

3. 거시적 의료비용 감소 여부

해당 결과 없음

4. 의료기관/의료인 효율 향상 여부

1) 자료 서지정보	Oktay O, Nanavati J, Schwaighofer A, Carter D, Bristow M, Tanno R, et al. Evaluation of Deep Learning to Augment Image-Guided Radiotherapy for Head and Neck and Prostate Cancers. JAMA Netw Open. 2020;3(11):e2027426.
2) 시험/검증 데이터/환자 기원	<ul style="list-style-type: none"> - 골반 CT: 2012년~2019년까지 5개 병원에서 전립선암으로 방사선치료를 받은 남자 환자 CT 총 132개 <ul style="list-style-type: none"> * 5개 병원 학습 데이터에서 무작위로 뽑은 91개(internal validation set 42 + test set 49) + 2개의 외부 병원(남미, 유럽) 83개(external validation set) - 두경부 CT: 2개 병원에서 두경부암으로 방사선치료를 받은 환자의 두경부 CT 66개 <ul style="list-style-type: none"> * 2개 병원 학습 데이터에서 무작위로 뽑은 40개(internal validation set 20 + test set 20) + 1개의 외부 병원(북미) 26개(external validation set)
3) 적용하려는 임상상황 및 대상환자와 부합하는가?	<input checked="" type="checkbox"/> 잘 부합함 <input type="checkbox"/> 불분명함 <input type="checkbox"/> 부합하지 않음(맞지 않는 환자군, 심한 selection bias)
4) 데이터/환자 수집 방법	<input checked="" type="checkbox"/> 후향적 <input type="checkbox"/> 전향적 <input type="checkbox"/> 전향적 및 후향적
5) 국가	<input type="checkbox"/> 한국데이터 <input checked="" type="checkbox"/> 외국데이터 <input type="checkbox"/> 한국 및 외국데이터: 한국 (%), 외국 (%), <input type="checkbox"/> 비율 불분명

<p>6) 결과 요약</p>	<p>핵심결과:</p> <p>평균 치료계획 시간 93% 단축</p> <p>1) 수동 분할:</p> <ul style="list-style-type: none"> - 두경부 CT: 평균 86.7분/scan (전문 reader), 평균 73.25분/scan (방사선종양학과 의사) <p>2) AI 사용 자동 분할 및 검토/교정:</p> <ul style="list-style-type: none"> - 두경부 CT: 평균 4.98분/scan - 골반 CT: 평균 3.40분/scan 	<p>AI 효과 판정</p> <p><input checked="" type="checkbox"/> 도움됨</p> <p><input type="checkbox"/> 불분명</p> <p><input type="checkbox"/> 도움 안됨</p>
-----------------	---	---

5. 환자중심결과 향상 측면

해당 결과 없음

예시 8

진단정확도 향상에 대한 불분명한 효과, 의료기관/의료인 효율 향상에 대한 불분명한 효과

예시 POINT

AI를 사용할 경우 AI를 사용하지 않는 경우에 비해 유방촬영을 이용한 유방암 검진에서 유방암 진단에 대한 정확도가 향상되는 일부 결과가 있으나, 이 결과를 이용해 추정해 보았을 때 net benefit이 있는지는 분명치 않음. 또한, 진료 효율에 대한 결과가 일관되지 않음.

가치 요소

- 의학적 효과: 진단정확도 향상에 대한 불분명한 효과
- 의료기관/의료인 효율 향상에 대한 불분명한 효과

I. AI 모델

1. AI 모델 요약

<p>1) 의료기기 인허가 자료 link</p>	<p>https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfpmn/pmn.cfm?ID=K211678 https://www.fda.gov/medical-devices/software-medical-device-samd/artificial-intelligence-and-machine-learning-aiml-enabled-medical-devices [Product Name] Computer Aided Detection / Diagnosis Software for Mammography / [Product License No.] 제 허 19-493호 / [Date of Initial Approval] 29 July 2019 / [Validity Period] 8 October 2020 ~ 31 May 2030 (식품의약품안전처)</p>
<p>2) 홈페이지 link</p>	<p>https://www.lunit.io/en/products/mmg</p>
<p>3) AI 모델의 기능/목적</p>	<p>Lunit INSIGHT MMG v1.1.1.0 - 기능: 유방촬영에서 유방 밀도 평가, 이상 병변을 감지하고 그 위치와 유방암 가능성 정도 점수 제시 - 목적: 여성 인구의 선별검진 유방촬영에서 유방암이 의심되는 영역을 감지하고, 위치 지정 및 특성화를 하여 의료인의 진단에 보조적인 도움을 주고 워크플로우 효율성을 증가시키기 위해 개발되었음.</p>
<p>4) 사용자를 위한 AI 출력 결과 형태</p>	<div data-bbox="491 1021 1018 1659" data-label="Image"> </div> <p>양측 상하위 촬영(Craniocaudal, CC view) 및 내외사위 촬영(mediolateral, MLO view)으로 총 4가지 사진에 대하여, 위치결정을 위한 heat map과 비정상 점수를 결과값으로 제시함.</p> <ol style="list-style-type: none"> 1) 위치결정(localization): 유방암이 의심되는 부위(종괴, 석회화, 모양 변형, 비대칭)에 heat map으로 표시 <ul style="list-style-type: none"> - 비정상 점수 역치값 10 이상인 곳을 표시 2) 비정상 점수(abnormality score) 제시: 0~100으로 평가점수를 부여하며, 100이 가장 높은 의심 수준.

<p>5) AI가 사용자에게 예측확률 또는 유사한 연속형 점수를 제시하는가?</p>	<p><input type="checkbox"/> 제시하지 않음</p> <p><input type="checkbox"/> 제시하며, 기준 값 선택 및 점수 해석에 대한 명확한 지침을 제공함</p> <p><input checked="" type="checkbox"/> 제시하며, 기준 값 선택 및 점수 해석에 대한 명확한 지침을 제공하지 않음</p>
--	--

2. 적용하려는 임상상황 및 대상환자

유방암에 대해 평균적인 위험을 가진 무증상 여성의 선별검진

3. 학습 데이터

<p>1) 학습 데이터 기원</p>	<p>3개의 한국 센터, 1개의 미국 센터, 1개의 영국 센터에서 모집된 170,230개의 선별검진 또는 진단 목적 모두를 포함한 유방촬영술 사진[조직학적으로 암으로 확진된 36,468개의 사진 및 양성으로 확인(조직검사 또는 최소 1년 추적촬영)된 59,544개의 사진 포함]</p>
<p>2) 적용하려는 임상상황 및 대상환자와 부합하는가?</p>	<p><input checked="" type="checkbox"/> 잘 부합함</p> <p><input type="checkbox"/> 불분명함</p> <p><input type="checkbox"/> 부합하지 않음</p>
<p>3) 국가</p>	<p><input type="checkbox"/> 한국데이터</p> <p><input type="checkbox"/> 외국데이터</p> <p><input checked="" type="checkbox"/> 한국 및 외국데이터: 한국 (85%), 외국 (15%), <input type="checkbox"/> 비율 불분명</p>

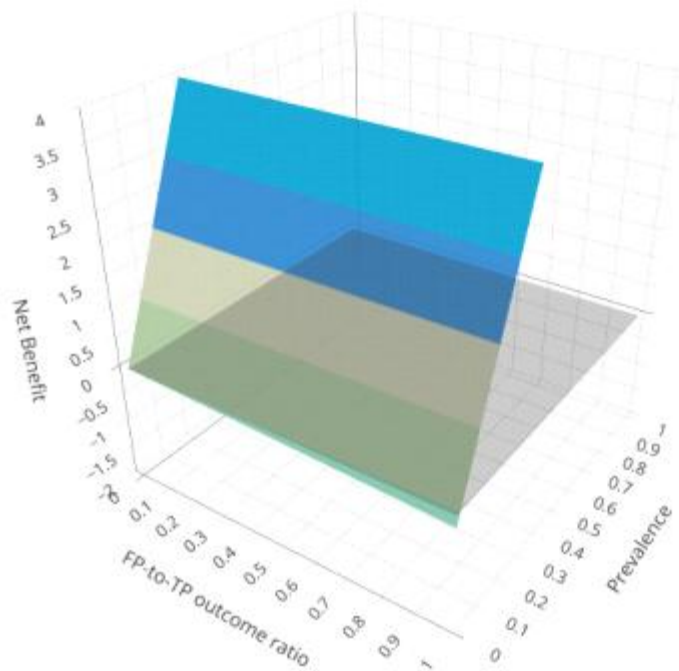
II. 임상 시험/검증 결과: 기존 진료와 AI를 활용한 진료 간 비교 결과

1. 의학적 효과 측면: 진단정확도 향상 여부 - 연구논문(1)

1) 자료 서지정보	Lee JH, Kim KH, Lee EH, Ahn JS, Ryu JK, Park YM, et al. Improving the Performance of Radiologists Using Artificial Intelligence-Based Detection Support Software for Mammography: A Multi-Reader Study. Korean J Radiol. 2022;23(5):505-516.
2) 시험/검증 데이터/환자 기원	조직학적으로 확진된 유방암 100례 및 유방암이 아닌 100례(양성 병변 40례, 정상 60례) 총 200개의 사례를 부천 순천향대병원에서 2009년 3월부터 2018년 9월까지 유방암 선별 검진으로 시행했던 유방촬영술로부터 모집하였음.
3) 적용하려는 임상상황 및 대상환자와 부합하는가?	<input checked="" type="checkbox"/> 잘 부합함 <input type="checkbox"/> 불분명함 <input type="checkbox"/> 부합하지 않음(맞지 않는 환자군, 심한 selection bias)
4) 데이터/환자 수집 방법	<input checked="" type="checkbox"/> 후향적 <input type="checkbox"/> 전향적 <input type="checkbox"/> 전향적 및 후향적
5) 국가	<input checked="" type="checkbox"/> 한국데이터 <input type="checkbox"/> 외국데이터 <input type="checkbox"/> 한국 및 외국데이터: 한국 (%), 외국 (%), <input type="checkbox"/> 비율 불분명
6) 결과 요약	<p>핵심결과: Study design: Paired design (Crossover) 5명의 유방전문영상의사와, 5명의 일반영상의로 구성된 총 10명의 영상의학전문목이 2 phase로 나누어 평가를 진행하였음. Phase 1에는 3인의 유방전문영상의사와 2인의 일반영상의가 AI 보조 하에 판독을, 나머지 2인의 유방전문영상의와 3인의 일반영상의는 AI 보조없이 판독을 하였음. 이후 2개월의 공백기를 가진 후 시행된 phase 2에서는 AI 보조를 받았던 의사는 AI 보조 없이, AI 보조가 없었던 의사는 AI 보조와 함께 판독을 시행하는 교차 연구를 진행함. 일차적 진단 성능은 전통적인 수신자 조작 특성 (ROC) 곡선 분석을 통해 제시됨.</p> <p>결과</p> <ul style="list-style-type: none"> ● AI 보조 판독 vs. AI 비보조 판독 (유방전문영상의 집단) <ol style="list-style-type: none"> 1) AUROC (area under the receiver operating characteristic curve) : AI 보조 시 비보조 보다 통계적으로 유의하게 상승(0.884 vs. 0.831, P = 0.007). 2) Sensitivity: AI 보조 시 비보조 보다 통계적으로 유의하게 상승 (88.6% vs. 74.6%, P < 0.001). 3) Specificity: AI 보조 시 비보조 보다 약간 감소하였으나, 통계적으로 유의한 차이 없음(66.4% vs. 66.6%, P = 0.238).

Net benefit (<https://aim-aicro.com/software/performanceoutcome>) 분석:

- Δ Sensitivity: (14%)
- Δ Specificity: (-0.2%)
- Disease prevalence: 0.3%~0.6% (출처: 2.5/1,000 [J Korean Soc Breast Screen 2004;1:15-20]; 6.2/1,000 [J Korean Soc Ultrasound Med 2010;29:1-5])
- Net benefit 분석결과: 암 검진을 위한 개별 유방촬영술 시행 때 유방암의 암 유병률이 매우 낮아 정확한 분석이 어려우며, net benefit이 있더라도 marginal한 benefit 예상

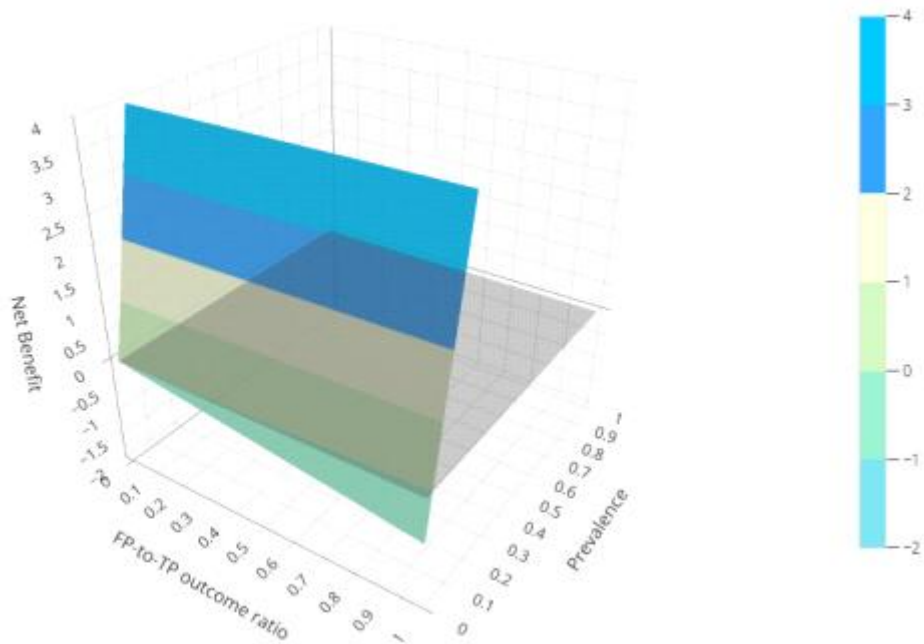


● AI 보조 판독 vs. AI 비보조 판독 (일반영상의 집단)

- 1) AUROC: AI 보조 시 비보조 보다 통계적으로 유의하게 상승(0.833 vs. 0.684, $P < 0.001$).
- 2) Sensitivity: AI 보조 시 비보조 보다 통계적으로 유의하게 상승(79.4% vs. 51.2%, $P < 0.001$).
- 3) Specificity: AI 보조 시 비보조 보다 약간 감소하였으나, 통계적으로 유의한 차이 없음(70.0% vs. 70.8%, $P = 0.689$).

Net benefit (<https://aim-aicro.com/software/performanceoutcome>) 분석:

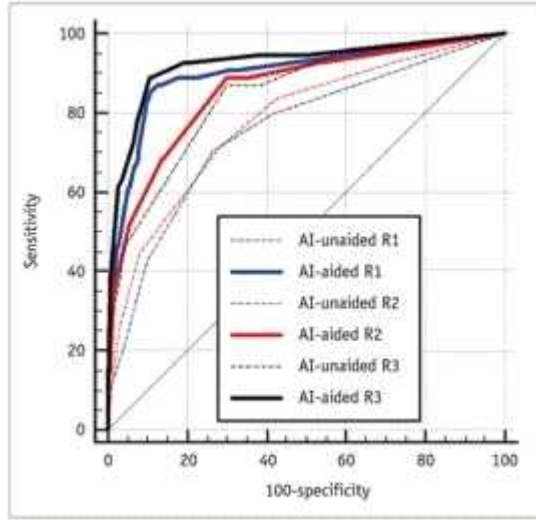
- Δ Sensitivity: (28.2%)
- Δ Specificity: (-0.8%)
- Disease prevalence: 0.3%~0.6% (출처: 2.5/1,000 [J Korean Soc Breast Screen 2004;1:15-20]; 6.2/1,000 [J Korean Soc Ultrasound Med 2010;29:1-5])
- Net benefit 분석결과: 암 검진을 위한 개별 유방촬영술 시행 때 유방암의 암 유병률이 매우 낮아 정확한 분석이 어려우며, net benefit이 있더라도 marginal한 benefit 예상



1. 의학적 효과 측면: 진단정확도 향상 여부 - 연구논문(2)

<p>1) 자료 서지정보</p>	<p>Kim YS, Jang MJ, Lee SH, Kim SY, Ha SM, Kwon BR, Moon WK, Chang JM. Use of Artificial Intelligence for Reducing Unnecessary Recalls at Screening Mammography: A Simulation Study. Korean J Radiol. 2022; 23(12):1241-1250.</p>	
<p>2) 시험/검증 데이터/환자 기원</p>	<p>서울대병원의 검진 센터 2곳에서 2016년 1월부터 2019년 12월 사이에 유방촬영술로 발견된 이상 소견에 대한 평가를 위해 추가 진단용 유방촬영술 또는 진단용 디지털 유방단층촬영술(digital breast tomosynthesis [DBT])를 받은 여성(=screening recall) 793명의 여성(30세~78세; 암 54명, 정상 혹은 양성 병변 739명)</p>	
<p>3) 적용하려는 임상상황 및 대상환자와 부합하는가?</p>	<p><input type="checkbox"/> 잘 부합함 <input type="checkbox"/> 불분명함 <input checked="" type="checkbox"/> 부합하지 않음(맞지 않는 환자군, 심한 selection bias)</p>	
<p>4) 데이터/환자 수집 방법</p>	<p><input checked="" type="checkbox"/> 후향적 <input type="checkbox"/> 전향적 <input type="checkbox"/> 전향적 및 후향적</p>	
<p>5) 국가</p>	<p><input checked="" type="checkbox"/> 한국데이터 <input type="checkbox"/> 외국데이터 <input type="checkbox"/> 한국 및 외국데이터: 한국 (%), 외국 (%), <input type="checkbox"/> 비율 불분명</p>	
<p>6) 결과 요약</p>	<p>핵심결과: 1. Study design: Paired design (Case-by case sequential reading) 하나의 사례에 대해 AI 없이 영상의학과 의사가 판독을 수행한 후, 동일한 사례에 대하여 AI 결과가 제공된 후 동일한 영상의학과 의사가 판독을 수행함.</p>	<p>AI 효과 판정 <input type="checkbox"/> 도움됨 <input checked="" type="checkbox"/> 불분명 <input type="checkbox"/> 도움 안됨</p>

2. AI 보조 대 비보조 결과 비교(n = 793, 95% 신뢰구간을 적용)



- 1) AUROC: AI 보조 시 비보조 보다 통계적으로 유의하게 상승 (0.89 vs. 0.79, $P < 0.001$).
- 2) Sensitivity: AI 보조 시 비보조 보다 상승하나, 통계적으로 유의하지 않음(92.6% vs. 89.5%, $P = 0.053$).
- 3) Specificity: AI 보조 시 비보조 보다 통계적으로 유의하게 상승 (53.9% vs. 41.9%, $P < 0.001$).
- 4) 재검율: AI 보조 시 비보조 보다 통계적으로 유의하게 감소 (49.5% vs. 60.4%, $P < 0.001$).

진단정확도 향상이 있어 보이나, 대상환자가 적용하려는 임상상황과 잘 부합하지 않아 결과 해석 및 적용에 있어 불분명한 제약이 있음.

2. 의학적 효과 측면: 진료/환자 결과 향상 여부

해당 결과 없음

3. 거시적 의료비용 감소 여부

해당 결과 없음

4. 의료기관/의료인 효율 향상 여부

1) 자료 서지정보	Lee JH, Kim KH, Lee EH, Ahn JS, Ryu JK, Park YM, et al. Improving the Performance of Radiologists Using Artificial Intelligence-Based Detection Support Software for Mammography: A Multi-Reader Study. Korean J Radiol. 2022;23(5):505-516.	
2) 시험/검증 데이터/환자 기원	조직학적으로 확진된 유방암 100례 및 유방암이 아닌 100례(양성 병변 40례, 정상 60례) 총 200개의 사례를 부천 순천향대병원에서 2009년 3월부터 2018년 9월까지 유방암 선별 검진으로 시행했던 유방촬영술로부터 모집하였음.	
3) 적용하려는 임상상황 및 대상환자와 부합하는가?	<input checked="" type="checkbox"/> 잘 부합함 <input type="checkbox"/> 불분명함 <input type="checkbox"/> 부합하지 않음(맞지 않는 환자군, 심한 selection bias)	
4) 데이터/환자 수집 방법	<input checked="" type="checkbox"/> 후향적 <input type="checkbox"/> 전향적 <input type="checkbox"/> 전향적 및 후향적	
5) 국가	<input checked="" type="checkbox"/> 한국데이터 <input type="checkbox"/> 외국데이터 <input type="checkbox"/> 한국 및 외국데이터: 한국 (%), 외국 (%), <input type="checkbox"/> 비율 불분명	
6) 결과 요약	<p>핵심결과: 5명의 유방전문영상의학과 5명의 일반영상의학과로 구성된 10명의 영상의학전문가가 2 phase로 나누어 평가를 진행하였음. 그룹 간에 AI 보조의 유무에 따른 판독시간을 비교함.</p> <p>AI 보조 판독 vs. AI 비보조 판독 시 판독시간 (유방전문영상의 집단)</p> <ol style="list-style-type: none"> 1) 전체: AI 보조 시 비보조 보다 통계적으로 유의하게 감소 (73.04초 vs. 82.73초, $P < 0.001$). 2) 유방암: AI 보조 시 비보조 보다 통계적으로 유의하게 감소 (74.79초 vs. 83.05초, $P = 0.004$). 3) 양성 질환: AI 보조 시 비보조 보다 통계적으로 유의하게 감소 (74.20초 vs. 86.41초, $P = 0.025$). 4) 정상: AI 보조 시 비보조 보다 통계적으로 유의하게 감소 (79.72초 vs. 69.37초, $P = 0.040$). <p>AI 보조 판독 vs. AI 비보조 판독 시 판독시간 (일반영상의 집단)</p> <ol style="list-style-type: none"> 1) 전체: AI 보조 시 비보조 보다 통계적으로 유의하게 증가 (42.52초 vs. 35.44초, $P < 0.001$). 2) 유방암: AI 보조 시 비보조 보다 통계적으로 유의하게 증가 (46.94초 vs. 37.99초, $P < 0.001$). 3) 양성 질환: AI 보조 시 비보조 보다 통계적으로 유의하게 증가 (47.45초 vs. 33.62초, $P < 0.001$). 4) 정상: AI 보조와 비보조 간에 통계적으로 유의미한 차이가 없음 ($P = 0.656$). <p>결론: AI 보조 시 유방전문영상의 집단에서는 모든 군에서 판독시간이 감소하였고, 일반영상의 집단에서는 유방암 및 양성 질환 판독에서 판독시간이 증가하였음.</p>	<p>AI 효과 판정</p> <input type="checkbox"/> 도움됨 <input checked="" type="checkbox"/> 불분명 <input type="checkbox"/> 도움 안됨

5. 환자중심결과 향상 측면

해당 결과 없음

예시 9

진단정확도 향상에 대한 불분명한 효과, 의료기관/의료인 효율 감소

예시 POINT


AI를 사용하였을 때 건강검진을 위한 흉부 방사선촬영 판독 진단정확도가 향상된 결과가 있으나 일관된 결과가 아니며, 이 결과를 이용해 진료/환자 결과에 있어 net benefit이 있는지 추정해 보았을 때 net benefit이 발생하는 조건이 제한적임. 또한, 의료인의 진료 효율은 오히려 감소하는 결과를 보여주고 있음.

가치 요소

- 의학적 효과: 진단정확도 향상에 대한 불분명한 효과
- 의료기관/의료인 효율 감소

I. AI 모델

1. AI 모델 요약

<p>1) 의료기기 인허가 자료 link</p>	<p>국내) Lunit INSIGHT CXR (루닛인사이트흉부단순방사선촬영), 2등급 의료 영상 검출 · 진단 보조 소프트웨어</p> <ul style="list-style-type: none"> - XR_KoreaMFDS: [품목명] 2등급의료영상검출·진단보조소프트웨어/ [품목허가번호] 제허 20-896호 / [최초허가일] 2020-10-19 / [유효기간] 2020-10-19 ~ 2025-10-18 [발급기관] 식품의약품안전처 - CXR_CE: Lunit의 (EU MDR) EC Certificate [First Issue Date] 2022-10-31/[Expiry Date] 2027-10-30 / [발급기관] BSI
<p>2) 홈페이지 link</p>	<p>https://www.lunit.io/ko/products/cxr</p>
<p>3) AI 모델의 기능/ 목적</p>	<p>Lunit INSIGHT CXR (루닛인사이트흉부단순방사선촬영)</p> <ul style="list-style-type: none"> - 기능: 흉부 단순방사선촬영에서 이상을 감지하여 대상 질병이 존재할 확률을 0에서 1 사이의 점수로 제시하고 병변의 위치를 heatmap으로 표기하여 시각화함. - 목적: 의료인의 진단 검출에 보조적 역할
<p>4) 사용자를 위한 AI 출력 결과 형태</p>	<p>병변 의심 부위 표시 및 heatmap, 이상 소견 점수를 제공함.</p>  <p>폐렴이 잘 감지된 사례의 예시, 왼쪽에서 오른쪽으로 순서대로 흉부 전산화단층촬영 영상, 흉부 단순방사선촬영 영상, 그리고 AI 소프트웨어에 의해 표기된 이상 병변의 위치와 heatmap, 비정상 점수</p> <p>참고로, 2023년 11월 기준 상용화 중인 LUNIT CXR은 흉부 단순방사선촬영에서 가장 흔하게 발견되는 10가지 비정상 소견(결절, 경화, 기흉, 흉막삼출, 무기폐, 기복증, 심장비대, 종격동확장, 석회화, 폐섬유화)을 검출하고 이상 병변의 검출 위치와 이상 부위 값을 표기하고 있음.</p>
<p>5) AI가 사용자에게 예측확률 또는 유사한 연속형 점수를 제시하는가?</p>	<p><input type="checkbox"/> 제시하지 않음</p> <p><input type="checkbox"/> 제시하며, 기준 값 선택 및 점수 해석에 대한 명확한 지침을 제공함</p> <p><input checked="" type="checkbox"/> 제시하며, 기준 값 선택 및 점수 해석에 대한 명확한 지침을 제공하지 않음</p>

2. 적용하려는 임상상황 및 대상환자

국내 일차의료기관 건강검진센터에 건강검진을 주소로 내원한 무증상 환자, 혹은 흉부 검진

3. 학습 데이터

1) 학습 데이터 기원	논문에 적용한 AI software (Lunit INSIGHT for CXR version 4.7.2)의 정확한 학습 데이터를 알 수 없음(학습 데이터가 계속 추가됨).
2) 적용하려는 임상상황 및 대상환자와 부합하는가?	<input type="checkbox"/> 잘 부합함 <input checked="" type="checkbox"/> 불분명함 <input type="checkbox"/> 부합하지 않음
3) 국가	<input type="checkbox"/> 한국데이터 <input type="checkbox"/> 외국데이터 <input type="checkbox"/> 한국 및 외국데이터: 한국 (%), 외국 (%), <input type="checkbox"/> 비율 불분명

II. 임상 시험/검증 결과: 기존 진료와 AI를 활용한 진료 간 비교 결과

1. 의학적 효과 측면: 진단정확도 향상 여부

1) 자료 서지정보	Kim C, Yang Z, Park SH, Hwang SH, Oh YW, Kang EY, et al. Multicentre external validation of a commercial artificial intelligence software to analyse chest radiographs in health screening environments with low disease prevalence. Eur Radiol. 2023;33(5):3501-3509.	
2) 시험/검증 데이터/환자 기원	국내 일차의료기관 건강검진센터 두 곳에서 2018년 1월부터 12월까지의 기간에 흉부 방사선촬영 및 전산화단층촬영을 동일한 날에 모두 실시한 총 3,047명의 무증상 환자(센터 A: 2,117명, 센터 B: 930명)	
3) 적용하려는 임상상황 및 대상환자와 부합하는가?	<input checked="" type="checkbox"/> 잘 부합함 <input type="checkbox"/> 불분명함 <input type="checkbox"/> 부합하지 않음(맞지 않는 환자군, 심한 selection bias)	
4) 데이터/환자 수집 방법	<input checked="" type="checkbox"/> 후향적 <input type="checkbox"/> 전향적 <input type="checkbox"/> 전향적 및 후향적	
5) 국가	<input checked="" type="checkbox"/> 한국데이터 <input type="checkbox"/> 외국데이터 <input type="checkbox"/> 한국 및 외국데이터: 한국 (%), 외국 (%), <input type="checkbox"/> 비율 불분명	
6) 결과 요약	<p>핵심결과:</p> <p>1. Study design: paired design, crossover</p> <ul style="list-style-type: none"> - 영상의학과 전공의 총 4명(1년차 2명, 3년차 2명)이 흉부 단순방사선촬영 영상을 2 session에 걸쳐 AI 보조/비보조 판독을 시행함(일차의료 검진 환경에서 CXR 해석 경험 수준이 적은 의사가 본 연구의 대상 제공자). - 판독 정확도를 평가하기 위해 AI 단독, 영상의학과 전공의의 AI 보조 유무 판독 각각 AUROC (area under the receiver operating characteristic curve), sensitivity, specificity를 계산하고 비교함. <p>2. 결과</p> <p>a) 임상적으로 중요한 질환으로 정의한 세 가지 병변(폐렴, 폐결핵, 폐암)의 유병률은 2.2% (68/3,047) 이었음.</p> <p>b) AI 보조 vs. AI 비보조</p> <ul style="list-style-type: none"> - AUROC: 판독의 4명 모두 AI 보조 시 통계적으로 유의하게 증가함. (0.571~0.688 vs. 0.534~0.676, 모든 P값 < 0.05) - Sensitivity: AI 보조(14.7%~44.1%) 및 AI 비보조(8.8%~41.2%)로 모든 판독의에서 증가했으나, 판독의 마다 통계적 유의성은 달랐음. 	<p>AI 효과 판정</p> <input type="checkbox"/> 도움됨 <input checked="" type="checkbox"/> 불분명 <input type="checkbox"/> 도움 안됨

- Specificity: AI 보조(93.5%~99.4%), AI 비보조(94.0%~99.3%)로 1명의 판독의를 제외한 3명의 판독의에서 AI 보조 시 감소함.

	AUROC (95% CI)		Sensitivity		Specificity	
	Without AI software assistance	With AI software assistance	Without AI software assistance	With AI software assistance	Without AI software assistance	With AI software assistance
Reader 1	0.676 (0.659-0.693)*	0.688 (0.671-0.705)†	41.2 (29.4-53.8)	44.1 (32.1-56.7)	94.0 (93.1-94.8)	93.5 (92.6-94.4)
Reader 2	0.534 (0.516-0.552)*	0.593 (0.575-0.611)†	8.8 (3.3-18.2)*	23.5 (14.1-35.4)†	98.0 (97.4-98.4)*	95.1 (94.3-95.8)†
Reader 3	0.553 (0.535-0.571)*	0.624 (0.606-0.641)†	11.8 (5.2-21.9)*	27.9 (17.7-40.1)†	98.8 (98.3-99.2)*	96.8 (96.1-97.4)†
Reader 4	0.563 (0.545-0.581)*	0.571 (0.553-0.588)†	13.2 (6.2-23.6)*	14.7 (7.3-25.4)	99.3 (99.0-99.6)*	99.4 (99.0-99.8)

AI software, artificial intelligence-based diagnostic assistant software; AUROC, area under the receiver operating curve; CI, confidence interval
 Note: Reader 1 [third-year]/Reader 2 [first-year] from Hospital 1; Reader 3 [third-year]/Reader 4 [first-year] from Hospital 2.
 **p* < 0.05 significance, reader vs AI software
 †*p* < 0.05 significance, with AI software assistance vs without AI software assistance

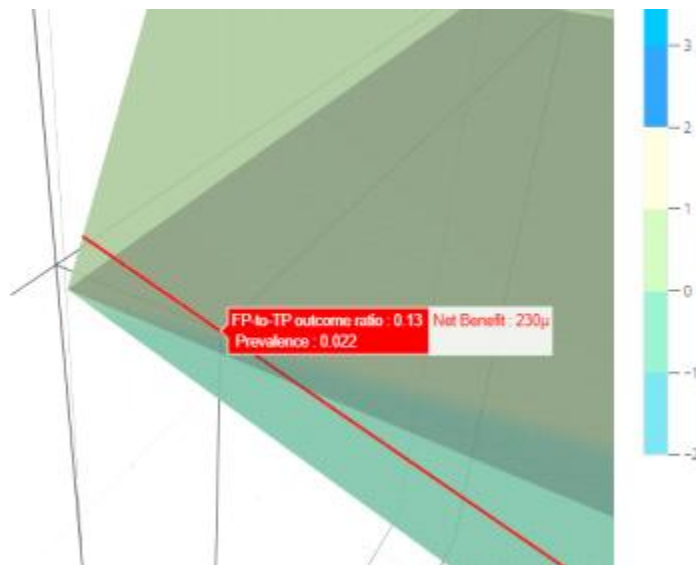
Net benefit (<https://aim-aicro.com/software/performance-to-outcome>) 분석:

가장 sensitivity가 높은 Reader 1의 결과를 적용하였을 때

- ΔSensitivity: (2.9%)
- ΔSpecificity: (-0.5%)
- Disease prevalence: 2.2%
- Net benefit 분석결과:

AI 사용에 따른 net benefit이 있으려면 FP-to-TP outcome ratio가 약 0.13 이하가 되어야 해서, net benefit이 발생하는 조건이 제한적임.

6) 결과 요약



2. 의학적 효과 측면: 진료/환자 결과 향상 여부

해당 결과 없음

3. 거시적 의료비용 감소 여부

해당 결과 없음

4. 의료기관/의료인 효율 향상 여부

1) 자료 서지정보	Kim C, Yang Z, Park SH, Hwang SH, Oh YW, Kang EY, et al. Multicentre external validation of a commercial artificial intelligence software to analyse chest radiographs in health screening environments with low disease prevalence. Eur Radiol. 2023;33(5):3501-3509.	
2) 시험/검증 데이터/환자 기원	국내 일차의료기관 건강검진센터 두 곳에서 2018년 1월부터 12월 기간에 흉부 방사선 촬영 및 전산화단층촬영을 동일한 날에 모두 실시한 총 3,047명의 무증상 환자(센터 A: 2,117명, 센터 B: 930명)	
3) 적용하려는 임상상황 및 대상환자와 부합하는가?	<input checked="" type="checkbox"/> 잘 부합함 <input type="checkbox"/> 불분명함 <input type="checkbox"/> 부합하지 않음(맞지 않는 환자군, 심한 selection bias)	
4) 데이터/환자 수집 방법	<input checked="" type="checkbox"/> 후향적 <input type="checkbox"/> 전향적 <input type="checkbox"/> 전향적 및 후향적	
5) 국가	<input checked="" type="checkbox"/> 한국데이터 <input type="checkbox"/> 외국데이터 <input type="checkbox"/> 한국 및 외국데이터: 한국 (%), 외국 (%), <input type="checkbox"/> 비율 불분명	
6) 결과 요약	<p>핵심결과: AI 보조 시 총 4명의 영상의학과 전공의의 흉부 단순방사선촬영 영상 판독시간은 AI 비보조와 비교하여 2.96초~10.27초 더 소요되었음(모든 P값 < 0.05).</p>	<p>AI 효과 판정</p> <input type="checkbox"/> 도움됨 <input type="checkbox"/> 불분명 <input checked="" type="checkbox"/> 도움 안됨

5. 환자중심결과 향상 측면

해당 결과 없음

예시 10

진단정확도 향상에 대한 불분명한 효과, 진료/환자 결과에 대한 불분명한 효과

예시 POINT

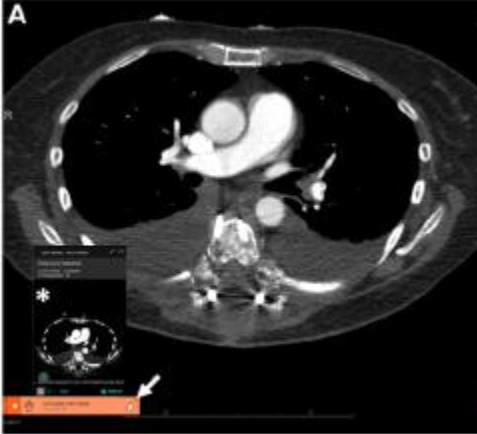
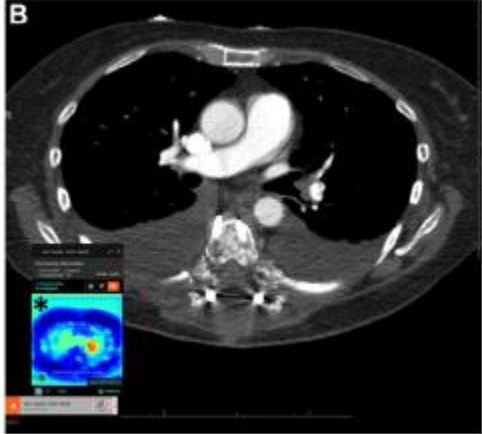
AI를 사용하였을 때 그렇지 않은 경우와 비교하여 폐색전증 의심 하에 폐혈관CT조영술을 시행 받은 환자의 영상 판독의 진단정확도 향상 및 환자 결과 향상이 불분명함

가치 요소

- 의학적 효과: 진단정확도 향상에 대한 불분명한 효과
- 의학적 효과: 진료/환자 결과에 대한 불분명한 효과

I. AI 모델

1. AI 모델 요약

<p>1) 의료기기 인허가 자료 link</p>	<p>Aidoc (version US 8.0 before November 2021 and US 8.1 after November 2021) https://medicalfuturist.com/fda-approved-ai-based-algorithms/ (FDA approval number: K180647 / type: 510(k)) https://grand-challenge.org/aiforradiology/product/aidoc-pulmonary-embolism/</p>
<p>2) 홈페이지 link</p>	<p>https://www.aidoc.com/solutions/radiology/</p>
<p>3) AI 모델의 기능/목적</p>	<ul style="list-style-type: none"> - 기능: 폐색전증 의심 하에 폐혈관CT조영술을 시행 받은 환자에 대하여 AI를 이용하여 폐색전증 여부를 선별함. 양성인 경우, 해당 사실을 영상의학과 의사에게 빠르게 알리고 의심되는 영상 소견을 제시함. - 목적: 폐색전증 양성으로 인지된 검사가 높은 우선순위와 높은 정확도로 판독되게 함.
<p>4) 사용자를 위한 AI 출력 결과 형태</p>	<p>폐색전증 양성으로 인지된 검사가 있을 때, 이를 영상의학과 의사에게 빠르게 알리며 의심되는 영상 소견을 제시함. 폐색전증 유무를 이분형으로 판단하고, 알람을 켜놓은 위젯이 열려있는 경우 양성으로 확인된 환자의 이름과 등록번호가 적혀있는 오렌지색의 직사각형 막대가 뜨면서 빠르게 알림.</p> <div style="display: flex; justify-content: space-around;">   </div>
<p>5) AI가 사용자에게 예측확률 또는 유사한 연속형 점수를 제시하는가?</p>	<ul style="list-style-type: none"> <input checked="" type="checkbox"/> 제시하지 않음 <input type="checkbox"/> 제시하며, 기준 값 선택 및 점수 해석에 대한 명확한 지침을 제공함 <input type="checkbox"/> 제시하며, 기준 값 선택 및 점수 해석에 대한 명확한 지침을 제공하지 않음

2. 적용하려는 임상상황 및 대상환자

폐색전증 의심 하에 폐혈관CT조영술을 시행 받은 환자

3. 학습 데이터

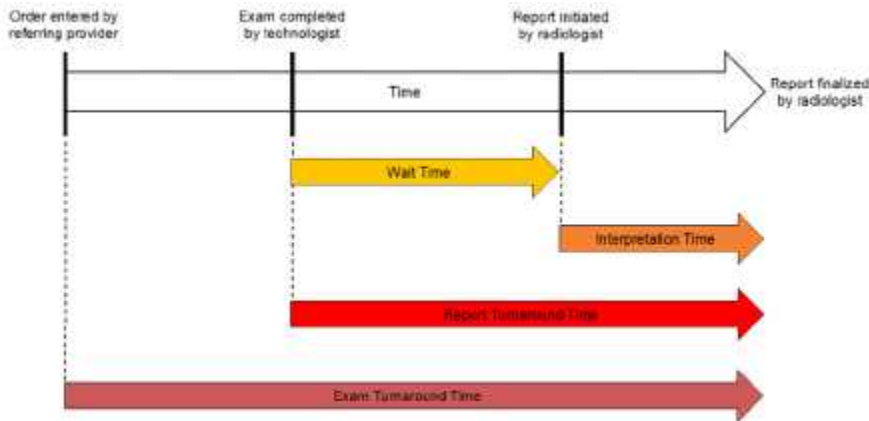
1) 학습 데이터 기원	알 수 없음
2) 적용하려는 임상상황 및 대상환자와 부합하는가?	<input type="checkbox"/> 잘 부합함 <input checked="" type="checkbox"/> 불분명함 <input type="checkbox"/> 부합하지 않음
3) 국가	<input type="checkbox"/> 한국데이터 <input type="checkbox"/> 외국데이터 <input type="checkbox"/> 한국 및 외국데이터: 한국 (%), 외국 (%), <input type="checkbox"/> 비율 불분명

II. 임상 시험/검증 결과: 기존 진료와 AI를 활용한 진료 간 비교 결과

1. 의학적 효과 측면: 진단정확도 향상 여부

1) 자료 서지정보	Rothenberg SA, Savage CH, AbouElkassam A, Singh S, Abozeed M, Hamki O, et al. Prospective Evaluation of AI Triage of Pulmonary Emboli on CT Pulmonary Angiograms. Radiology. 2023;309(1):e230702.	
2) 시험/검증 데이터/환자 기원	미국 University of Alabama 소속의 3차 병원에서 2021년 5월 12일 ~ 2021년 12월 4일 사이에 폐색전증 의심 하에 시행된 1,526건의 폐혈관CT조영술(1,436명의 환자)	
3) 적용하려는 임상상황 및 대상환자와 부합하는가?	<input checked="" type="checkbox"/> 잘 부합함 <input type="checkbox"/> 불분명함 <input type="checkbox"/> 부합하지 않음(맞지 않는 환자군, 심한 selection bias)	
4) 데이터/환자 수집 방법	<input type="checkbox"/> 후향적 <input checked="" type="checkbox"/> 전향적 <input type="checkbox"/> 전향적 및 후향적	
5) 국가	<input type="checkbox"/> 한국데이터 <input checked="" type="checkbox"/> 외국데이터 <input type="checkbox"/> 한국 및 외국데이터: 한국 (%), 외국 (%), <input type="checkbox"/> 비율 불분명	
6) 결과 요약	<p>핵심결과: 폐색전증 의심 하에 시행된 1,526건의 폐혈관CT조영술 및 1,448명의 환자를 대상으로 연구를 진행함.</p> <p>(가) AI 보조 vs. AI 비보조</p> <p>(a) 폐색전증 검출 정확도와 누락율(miss rate)은 AI 비보조 시 각각 97.6%와 12.3%, AI 보조 시 각각 98.6%와 6.1%로 AI 보조 시 정확도가 약간 높았으나, 통계적으로 유의한 차이가 없었음(각각 P = 0.15, P = 0.11).</p> <p>(b) 임상적으로 유의하지 않은 폐색전증에 대해서는 AI 보조 시 sensitivity가 통계적으로 유의하게 증가하였음(86.4% vs. 66.7%, P= 0.04).</p>	<p>AI 효과 판정</p> <input type="checkbox"/> 도움됨 <input checked="" type="checkbox"/> 불분명 <input type="checkbox"/> 도움 안됨

2. 의학적 효과 측면: 진료/환자 결과 향상 여부

1) 자료 서지정보	Rothenberg SA, Savage CH, AbouElkassam A, Singh S, Abozeed M, Hamki O, et al. Prospective Evaluation of AI Triage of Pulmonary Emboli on CT Pulmonary Angiograms. Radiology. 2023;309(1):e230702.
2) 시험/검증 데이터/환자 기원	미국 University of Alabama 소속의 3차 병원에서 2021년 5월 12일 ~ 2021년 12월 4일 사이에 폐색전증이 의심되어 시행된 1,526건의 폐혈관CT조영술(1,436명의 환자)
3) 적용하려는 임상상황 및 대상환자와 부합하는가?	<input checked="" type="checkbox"/> 잘 부합함 <input type="checkbox"/> 불분명함 <input type="checkbox"/> 부합하지 않음(맞지 않는 환자군, 심한 selection bias)
4) 데이터/환자 수집 방법	<input type="checkbox"/> 후향적 <input checked="" type="checkbox"/> 전향적 <input type="checkbox"/> 전향적 및 후향적
5) 국가	<input type="checkbox"/> 한국데이터 <input checked="" type="checkbox"/> 외국데이터 <input type="checkbox"/> 한국 및 외국데이터: 한국 (%), 외국 (%), <input type="checkbox"/> 비율 불분명
6) 결과 요약	<p>핵심결과:</p> <ul style="list-style-type: none"> - Phase 1 (2021년 5월 12일 ~ 2021년 6월 30일)과 phase 2 (2021년 9월 30일 ~ 2021년 12월 5일)로 기간을 나누어, 해당 시기에 폐색전증 의심 하에 시행된 폐혈관CT조영술(각 503건/489명, 1,023건/959명의 검사)에 대하여 연구를 진행함. - Phase 1에서는 영상의학과 의사가 기존의 판독업무 진행 방식으로 판독, Phase 2에서는 AI의 기능을 활용하는 환경에서 판독업무를 진행하였음. 둘 간에 판독을 시작할 때까지 및 판독 결과가 나올 때까지 소요되는 시간(wait time, report turnaround time, exam turnaround time)을 비교하였음.  <p>AI 효과 판정</p> <input type="checkbox"/> 도움됨 <input checked="" type="checkbox"/> 불분명 <input type="checkbox"/> 도움 안됨
<p>- 폐색전증 양성 시 평균 wait time(검사 종료시점부터 판독 시작 시점까지의 시간)은 AI 보조 시 AI 비보조 보다 통계적으로 유의하게 짧았음(11.3분 vs. 21.5분, P <0.001).</p>	

	<ul style="list-style-type: none"> - 폐색전증 양성 시 exam turnaround time(검사 처방 시점부터 판독 확정 시점까지의 시간)과 report turnaround time(검사 종료 시점부터 판독 확정 시점까지의 시간)은 AI 보조와 비보조 간에 통계적으로 유의한 차이가 없었음. - 모든 검사의 평균 report turnaround time(검사 종료 시점부터 판독 확정 시점까지의 시간)은 AI 보조 시 비보조 보다 통계적으로 유의미하게 짧았음 (45.0분 vs. 53.7분, $P < 0.001$). - 이러한 시간의 단축이 진료/환자 결과에 실제로 미치는 영향은 확인되지 않았음. 	
--	--	--

3. 거시적 의료비용 감소 여부

해당 결과 없음

4. 의료기관/의료인 효율 향상 여부

해당 결과 없음

5. 환자중심결과 향상 측면

해당 결과 없음

예시 11

환자중심결과에 대한 불분명한 효과

예시 POINT

AI를 사용할 경우 그렇지 않은 경우에 비해 전이성 암 환자의 말기 환자 진료에 있어 환자중심결과의 개선 가능성을 보여주었으나, 명확하지는 않은 예

가치 요소

- 환자중심결과에 대한 불분명한 효과

I. AI 모델

1. AI 모델 요약

1) 의료기기 인허가 자료 link	인허가를 받은 AI 모델이 아님
2) 홈페이지 link	
3) AI 모델의 기능/목적	전이암 환자를 대상으로 EMR 기록을 입력하면 기대여명을 예측하는 예후 예측 모델(prognostic model) 개발을 통해 종양내과 의사들이 환자의 예후를 예측하는데 도움.
4) 사용자를 위한 AI 출력 결과 형태	환자의 예측 생존율 곡선(survival curve), 중위 생존기간이 제공됨.
5) AI가 사용자에게 예측확률 또는 유사한 연속형 점수를 제시하는가?	<input type="checkbox"/> 제시하지 않음 <input type="checkbox"/> 제시하며, 기준 값 선택 및 점수 해석에 대한 명확한 지침을 제공함 <input checked="" type="checkbox"/> 제시하며, 기준 값 선택 및 점수 해석에 대한 명확한 지침을 제공하지 않음

2. 적용하려는 임상상황 및 대상환자

전이성 암 환자 중 예상 수명이 짧은 고위험 환자를 선별, ACP 대화를 위한 환자 선정을 최적화하여 ACP 문서 작성 빈도를 증가시키고 치료 강도를 감소시켜 최종적으로 환자 삶의 질 향상이 목표.

*ACP는 Advance Care Planning의 약자로, 말기 암 환자에서 치료자와 환자 간의 대화를 통해 치료계획을 수립하는 것을 의미함.

3. 학습 데이터

1) 학습 데이터 기원	- 2008년부터 2017년까지 미국 Stanford Health Care System에서 전이성 암으로 치료받은 10,070명. - 환자의 의무기록지, 활력 징후, lab 검사 결과, 시술, 처방, 진단 코드가 포함됨. https://doi.org/10.1093/jnci/djy178
2) 적용하려는 임상상황 및 대상환자와 부합하는가?	<input checked="" type="checkbox"/> 잘 부합함 <input type="checkbox"/> 불분명함 <input type="checkbox"/> 부합하지 않음
3) 국가	<input type="checkbox"/> 한국데이터 <input checked="" type="checkbox"/> 외국데이터 <input type="checkbox"/> 한국 및 외국데이터: 한국 (%), 외국 (%), <input type="checkbox"/> 비율 불분명

II. 임상 시험/검증 결과: 기존 진료와 AI를 활용한 진료 간 비교 결과

1. 의학적 효과 측면: 진단정확도 향상 여부

해당 결과 없음

2. 의학적 효과 측면: 진료/환자 결과 향상 여부

해당 결과 없음

3. 거시적 의료비용 감소 여부

해당 결과 없음

4. 의료기관/의료인 효율 향상 여부

해당 결과 없음

5. 환자중심결과 향상 측면

1) 자료 서지정보	Gensheimer MF, Gupta D, Patel MI, Fardeen T, Hildebrand R, Teuteberg W, et al. Use of Machine Learning and Lay Care Coaches to Increase Advance Care Planning Conversations for Patients With Metastatic Cancer. JCO Oncol Pract. 2023;19(2):e176-e184.
2) 시험/검증 데이터/환자 기원	미국 Stanford Health Care의 종양내과 클리닉 4곳의 총 1,251명 환자
3) 적용하려는 임상상황 및 대상환자와 부합하는가?	<input checked="" type="checkbox"/> 잘 부합함 <input type="checkbox"/> 불분명함 <input type="checkbox"/> 부합하지 않음(맞지 않는 환자군, 심한 selection bias)
4) 데이터/환자 수집 방법	<input type="checkbox"/> 후향적 <input checked="" type="checkbox"/> 전향적 <input type="checkbox"/> 전향적 및 후향적
5) 국가	<input type="checkbox"/> 한국데이터 <input checked="" type="checkbox"/> 외국데이터 <input type="checkbox"/> 한국 및 외국데이터: 한국 (%), 외국 (%), <input type="checkbox"/> 비율 불분명

6) 결과 요약	<p>핵심결과: Study design: parallel design 1) 흉부/비뇨기과 종양 클리닉 환자는 실험군, 피부/육종 종양 클리닉 환자는 대조군으로 사용 2) 실험군에서는 매주 AI로 예측한 기대여명 결과(2년 이내인 환자는 별도로 강조)를 이메일로 받고, care coach (종양학, 완화치료 및 ACP에 대해 훈련 받은 비의사 직종)가 이러한 환자군에게 ACP 대화를 시행 및 기록 3) 개입이 끝난 시점에서 ACP 문서 작성율, 예후 문서 작성율, end-of-life 치료 강도를 비교</p> <p>결과 1) ACP 문서 작성율, 예후 문서 작성율</p> <table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th rowspan="2"></th> <th colspan="2">실험군</th> <th colspan="2">대조군</th> </tr> <tr> <th>개입 전</th> <th>개입 후</th> <th>개입 전</th> <th>개입 후</th> </tr> </thead> <tbody> <tr> <td>ACP 문서 작성율</td> <td style="text-align: center;">3% (10/358)</td> <td style="text-align: center;">35% (315/910)</td> <td style="text-align: center;">3% (9/111)</td> <td style="text-align: center;">3% (9/341)</td> </tr> <tr> <td>예후 문서 작성율</td> <td style="text-align: center;">2% (7/358)</td> <td style="text-align: center;">27% (243/910)</td> <td style="text-align: center;">0% (0/111)</td> <td style="text-align: center;">1% (5/341)</td> </tr> </tbody> </table> <p>2) End-of-life 치료 강도: 평균적으로는 실험군과 대조군에서 end-of-life 치료 강도가 유사했지만, 적어도 1명의 제공자 ACP 편집이 있는 환자는 고강도 치료 조치가 더 적음(P = 0.04).</p> <p>이러한 차이가 환자의 삶의 질 향상과 불안 감소에 영향을 주는지에 대한 뚜렷한 결과는 제시되지 않음.</p> <p>Care coach를 통한 예측 생존기간의 제공은 ACP의 문서화와는 유의한 상관관계, 고강도 치료 조치와 음의상관관계가 확인되었으나, 환자중심결과 (patient-centered outcome [POC], 본문에서 POC로 언급되었던 삶의 질 또는 goal-concordant care)에 대한 뚜렷한 결과는 제시되지 않음.</p>					실험군		대조군		개입 전	개입 후	개입 전	개입 후	ACP 문서 작성율	3% (10/358)	35% (315/910)	3% (9/111)	3% (9/341)	예후 문서 작성율	2% (7/358)	27% (243/910)	0% (0/111)	1% (5/341)
		실험군		대조군																			
개입 전		개입 후	개입 전	개입 후																			
ACP 문서 작성율	3% (10/358)	35% (315/910)	3% (9/111)	3% (9/341)																			
예후 문서 작성율	2% (7/358)	27% (243/910)	0% (0/111)	1% (5/341)																			
<p>AI 효과 판정</p> <p><input type="checkbox"/> 도움됨</p> <p><input checked="" type="checkbox"/> 불분명</p> <p><input type="checkbox"/> 도움 안됨</p>																							



발행일 2024년 5월
발행처 한국보건의료연구원
주소 (04933) 서울특별시 광진구 능동로 400 (중곡동)보건복지행정타운 3~5층
대표전화번호 02-2174-2700
홈페이지 <https://www.neca.re.kr/>
ISBN 979-11-93112-86-1