THE POWER OF OPPORTUNISTIC FINDINGS USING AI

THE VALUE PERSPECTIVE

Dr. Jacob J. Visser, radiologist, health economist, clinical epidemiologist

Assistant professor Value-based imaging

CMIO

Erasmus MC

DISCLOSURES

- Medical advisor Noaber Foundation
- Medical advisor NLC
- Medical advisor Contextflow
- Medical advisor Quibim
- Research / travel grant Qure.ai

IMPACT

- Overall opinion medical imaging AI is positive
- High proportion of radiologists believe in positive impact
- Strategic AI-enabled could help in early intervention by getting more patients on the treatment pathway sooner (e.g. heart failure, lung cancer, COPD)

Original Article | Open Access | Published: 25 April 2022

An evaluation of information online on artificial intelligence in medical imaging

Philip Mulryan, Naomi Ni Chleirigh, Alexander T. O'Mahony, Claire Crowley, David Ryan, Patrick McLaughlin, Mark McEntee, Michael Maher & Owen J. O'Connor

Insights into Imaging **13**, Article number: 79 (2022) | Cite this article

Erasmus MC

EVALUATION

MOCK SCORING RUBRIC

Local Performance Metrics/ Ease of Use

0 – Poor local performance, cumbersome UX
10 – Modest local performance, useable UX
20 – Good local performance, good UX
30 – Excellent local performance, excellent UX

Technical Readiness / Workflow Impact

0 – No existing technical infrastructure
10 – Major technical modification
20 – Minor technical modification
30 – All technical infrastructure in place

Scientific Evidence

0 – none / pilot data
10 – Level III/IV limited cohort
20 – Level II prospective clinical evaluation
30 – Level I evidence for primary outcome

Value

0 – Costs unknown
5 – Net negative
10 – Net neutral
15 – Billing code

Fairness/Bias/Harm

0 – unknown
10 – High
20 – Moderate
30 – Lowest risk

Clinical Impact

0 – Low volume low acuity
10 – Moderate volume or moderate acuity
20 – High volume or high acuity
30 – High volume and high acuity

Daye et al, Implementation of Clinical Artificial Intelligence in Radiology: Who Decides and How?; Radiology August 2022
https://pubs.rsna.org/doi/full/10.1148/radiol.212151

EVALUATION

Scientific evidence

- Standard metrics include:
  - Area under the receiver operating characteristic curve; sensitivity and specificity; the precision-recall curve; and, when applicable, regression metrics (root mean square error, mean absolute error, R^2)

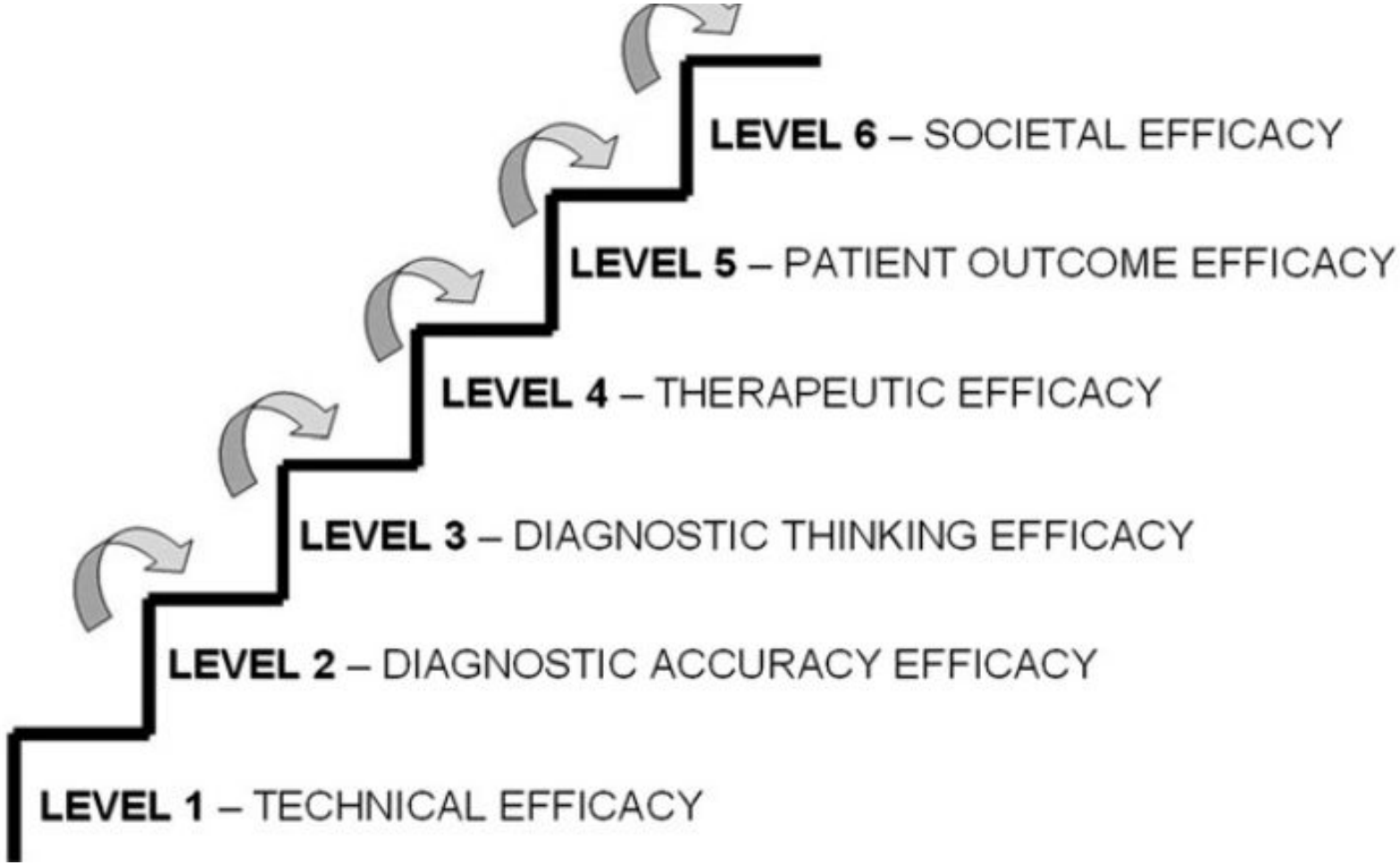
- Guidelines have been published for evaluating and reporting the results of AI models:
  - Transparent Reporting of a Multivariable Prediction Model of Individual Prognosis or Diagnosis (TRIPOD-AI)
  - Standards for Reporting of Diagnostic Accuracy Studies (known as STARD-AI)
  - Consolidation Standards of Reporting Trials (known as CONSORT-AI)

Erasmus MC

EVALUATION

Scientific evidence

- Fryback model

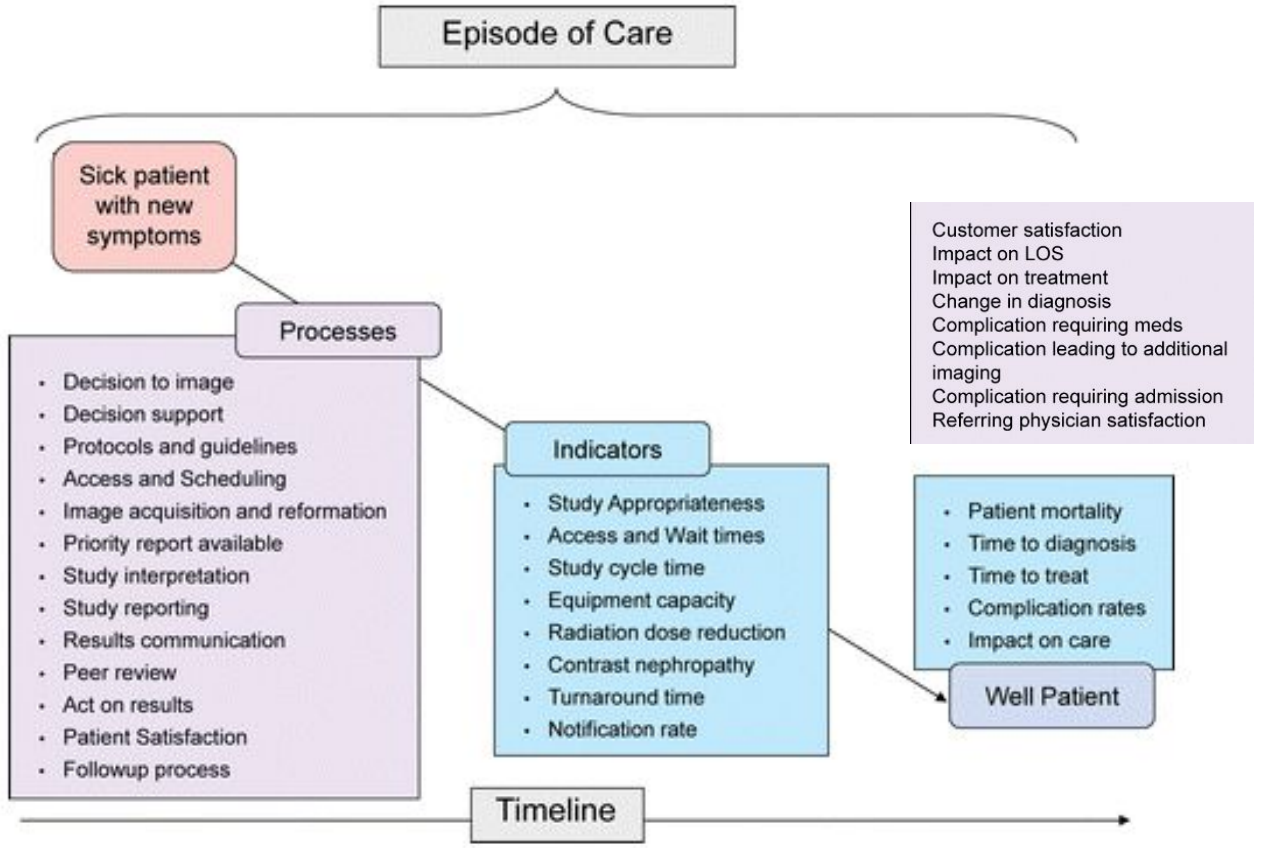


Fryback DG, Thornbury JR. The efficacy of diagnostic imaging. Med Decis Making. 1991 Apr-Jun;11(2):88-94. doi: 10.1177/0272989X9101100203. PMID: 1907710.

EVALUATION

Clinical impact

- Distinguish between:
  - Process metrics
  - Outcomes metrics
  - Value metrics

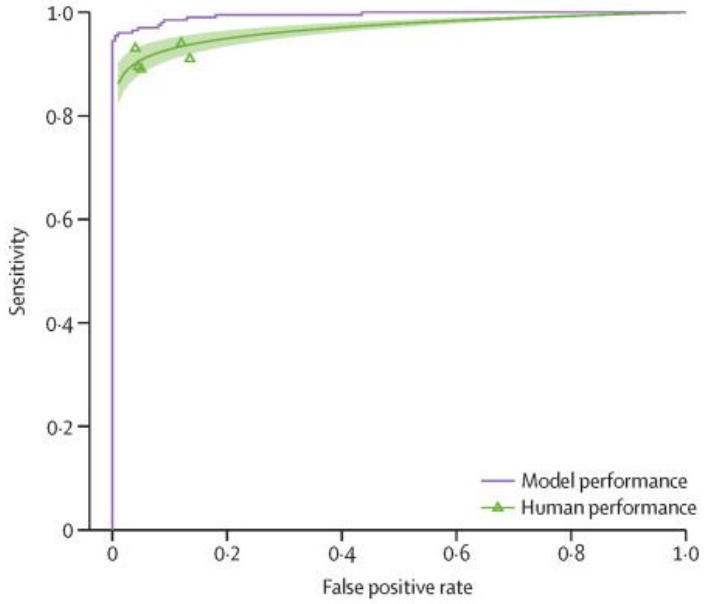


Sarwar A, Boland G, Monks A, Kruskal JB. Metrics for Radiologists in the Era of Value-based Health Care Delivery. Radiographics. 2015 May-Jun;35(3):866-76. doi: 10.1148/rg.2015140221. Epub 2015 Apr 3. PMID: 25839737.

EVALUATION

Local situation

- Most deep learning algorithms show fall in accuracy in external datasets
- Lack of diverse datasets in AI research puts patients at risk
  - Algorithm detecting proximal femoral fractures
  - Model outperformed humans
    - Operating point needed to change in external validation set
    - Unexpected and potentially harmful algorithm behavior (abnormal bones)



Oakden-Rayner L, Gale W, Bonham TA, Lungren MP, Carneiro G, Bradley AP, Palmer LJ. Validation and algorithmic audit of a deep learning system for the detection of proximal femoral fractures in patients in the emergency department: a diagnostic accuracy study. Lancet Digit Health. 2022 May;4(5):e351-e358. doi: 10.1016/S2589-7500(22)00004-8. Epub 2022 Apr 5. PMID: 35396184.

Erasmus MC

EVALUATION

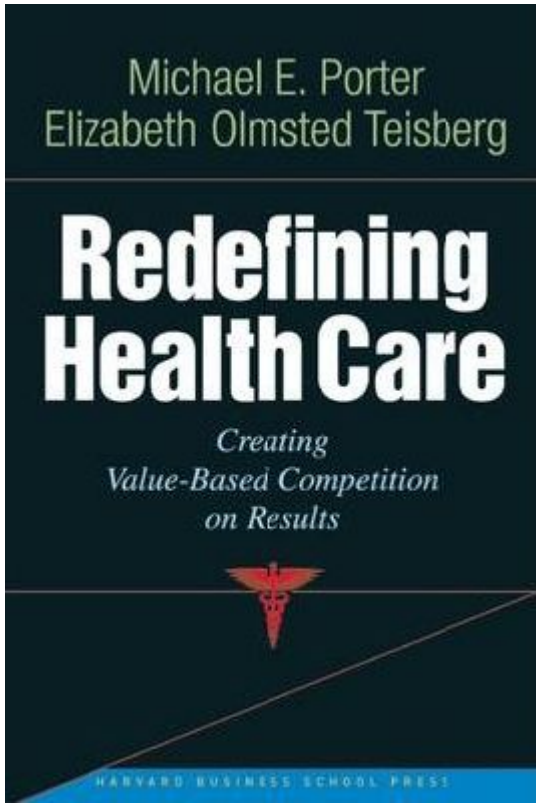
Local situation

- Data set used for model development congruent with setting in which model will be used
  - In- and exclusion criteria applied
- Unbiased local dataset:
  - Accuracy and reproducability
    - > Sensitivity / specificity versus clinical utility thresholds
- Identify effects along health care chain
- Technical / workflow impact

Trivedi H. The Business of Artificial Intelligence in Radiology Has Little to Do With Radiologists. J Am Coll Radiol. 2022 Apr;19(4):564-566. doi: 10.1016/j.jacr.2022.01.006. Epub 2022 Feb 15. PMID: 35176248.

EVALUATION

Value



$$\text{Patient Value} = \frac{\text{Health Outcomes}}{\text{Cost}}$$

Erasmus MC

EVALUATION

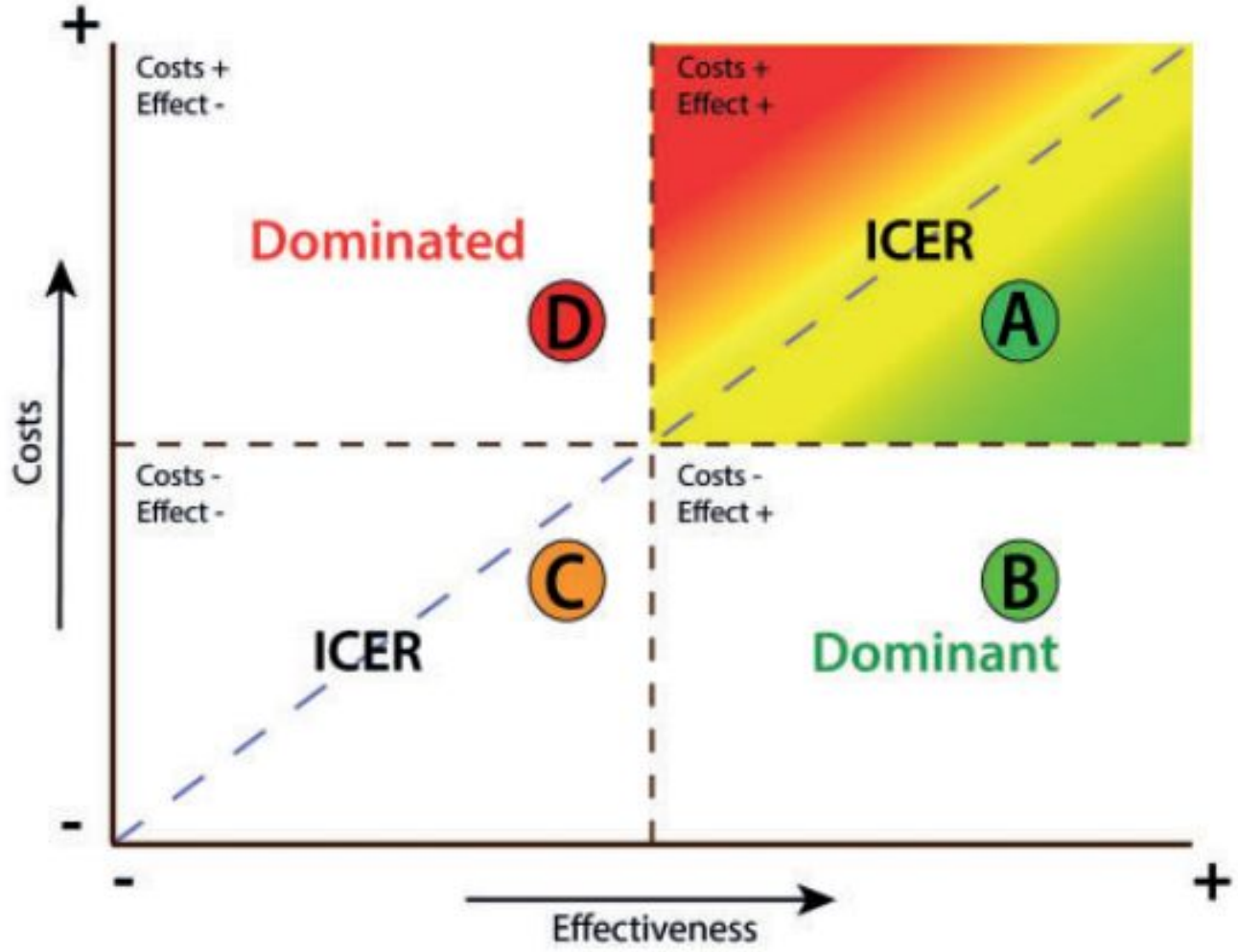
Value

Cost-effectiveness analysis	Value-based healthcare
$ICER = \frac{Costs\ (A\text{-}B)}{Outcomes\ (A\text{-}B)}$	$Value = \frac{Outcomes}{Costs}$

ICER indicates incremental cost-effectiveness ratio.

EVALUATION

Value



Antonides CFJ, Cohen DJ, Osnabrugge RLJ. Statistical primer: a cost-effectiveness analysis. Eur J Cardiothorac Surg. 2018 Aug 1;54(2):209-213. doi: 10.1093/ejcts/ezy187. PMID: 29726940.

Erasmus MC

CONCLUSIONS

- Scientific evidence
  - Standard metrics
  - 6-level Fryback model
- Clinical impact
  - Process vs outcome vs value metrics
- Thorough evaluation needed of AI-algorithms before implementation
  - Local performance metrics
    - Hospital / region / country
    - Consider whole health care chain
- Value
  - Need to consider costs along health care chain
  - Combine cost-effectiveness analysis and value-based healthcare


Erasmus MC
Erasmus MC Hoofdingang
Spoedeisende hulp