

TECNOLOGIE AVANZATE ED AI PER MIGLIORARE
I PERCORSI DIAGNOSTICI E DI CURA:
Buone pratiche nelle aziende sanitarie

Criticità nella diagnostica per immagini: l'importanza di un programma di garanzia di sicurezza e qualità

CARLO CAVEDON

UOC Fisica Sanitaria
Dipartimento di Patologia e
Diagnostica
Azienda Ospedaliera
Universitaria Integrata - Verona



1. detection / ausilio alla diagnosi

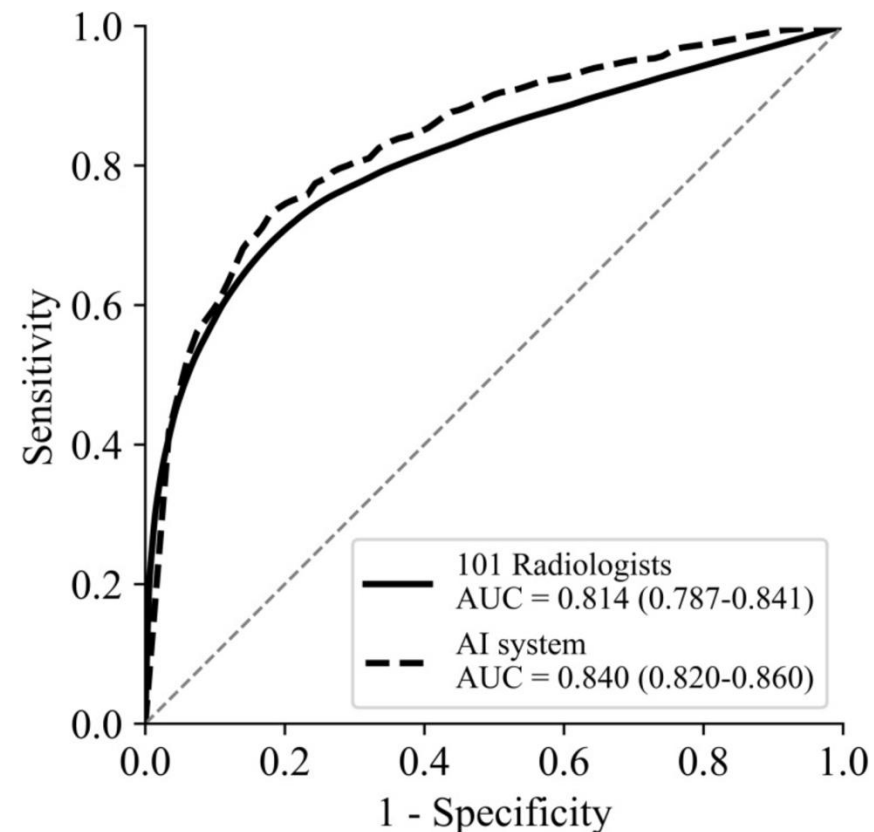
> J Natl Cancer Inst. 2019 Sep 1;111(9):916-922. doi: 10.1093/jnci/djy222.

Stand-Alone Artificial Intelligence for Breast Cancer Detection in Mammography: Comparison With 101 Radiologists

Alejandro Rodriguez-Ruiz, Kristina Lång, Albert Gubern-Merida, Mireille Broeders, Gisella Gennaro, Paola Clauser, Thomas H Helbich, Margarita Chevalier, Tao Tan, Thomas Mertelmeier, Matthew G Wallis, Ingvar Andersson, Sophia Zackrisson, Ritse M Mann, Ioannis Sechopoulos

PMID: 30834436 PMCID: [PMC6748773](#) DOI: [10.1093/jnci/djy222](#)

Results: The performance of the AI system was statistically noninferior to that of the average of the 101 radiologists. The AI system had a 0.840 (95% confidence interval [CI] = 0.820 to 0.860) area under the ROC curve and the average of the radiologists was 0.814 (95% CI = 0.787 to 0.841) (difference 95% CI = -0.003 to 0.055). The AI system had an AUC higher than 61.4% of the radiologists.

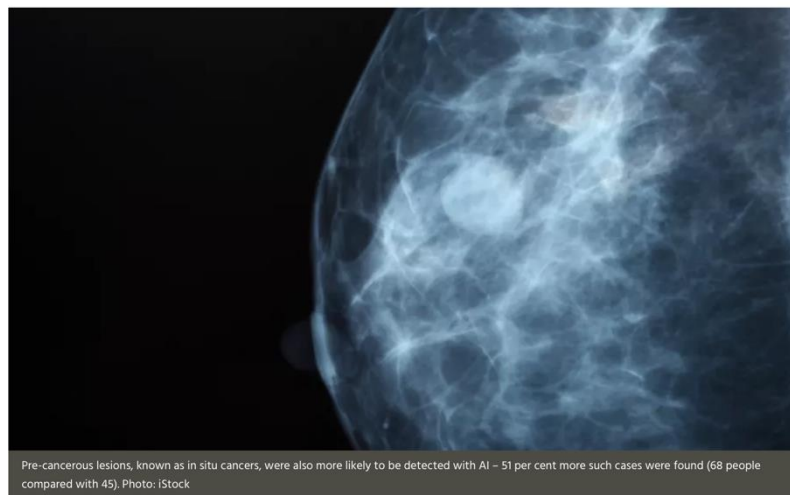


1. detection / ausilio alla diagnosi: AI come *second reader*

[Start](#) > [News](#) > AI-supported breast cancer screening – new results suggest even higher accuracy

AI-supported breast cancer screening – new results suggest even higher accuracy

By Erika Svantesson - published 4 February 2025



New research results now published from Lund University's MASAI trial are even better than the initial findings from last year: AI-supported breast screening detected 29 per cent more cases of cancer compared with traditional screening. More invasive cancers were also clearly detected at an early stage using AI. Now the final part of the research study will focus on breast cancer missed by screening.

Contact

Kristina Lång, researcher and associate professor of diagnostic radiology at Lund University, Sweden and consultant at the Unilabs Mammography Unit in Malmö.

Randomized Controlled Trial > [Lancet Digit Health](#). 2025 Mar;7(3):e175-e183.

doi: 10.1016/S2589-7500(24)00267-X. Epub 2025 Feb 3.

Screening performance and characteristics of breast cancer detected in the Mammography Screening with Artificial Intelligence trial (MASAI): a randomised, controlled, parallel-group, non-inferiority, single-blinded, screening accuracy study

Veronica Hernström¹, Viktoria Josefsson¹, Hanna Sartor², David Schmidt², Anna-Maria Larsson³, Solveig Hofvind⁴, Ingvar Andersson², Aldana Rosso⁵, Oskar Hagberg⁵, Kristina Lång⁶

- MASAI - ca. 106k casi
- soglia dimostrazione sicurezza: 3/1000
- *risultati: 338/53k AI assisted (6.4/1000) vs. 262/53k (4.9/1000) standard doppia lettura*
- no aumento del rateo di falsi positivi

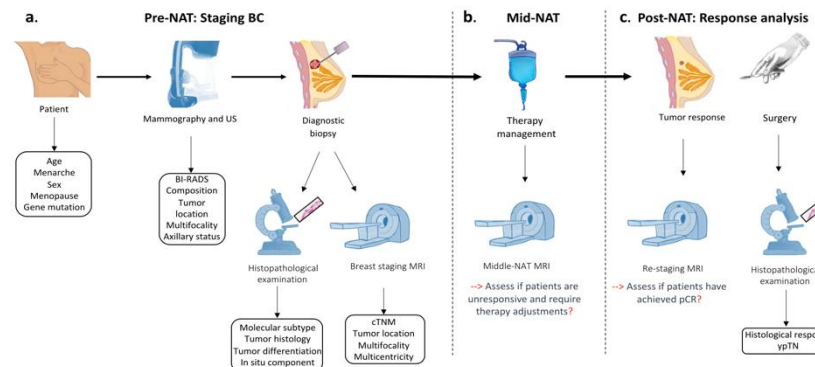
2. modelli predittivi

Multicenter Study > Nat Commun. 2024 Nov 7;15(1):9613.

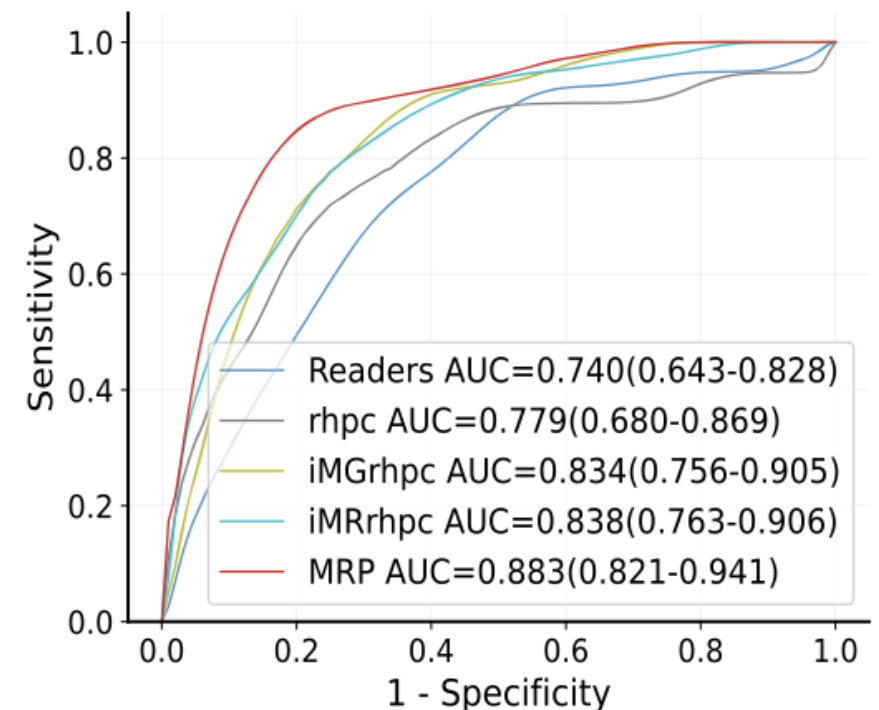
doi: 10.1038/s41467-024-53450-8.

An explainable longitudinal multi-modal fusion model for predicting neoadjuvant therapy response in women with breast cancer

Yuan Gao^{1 2 3}, Sofia Ventura-Diaz⁴, Xin Wang^{1 2 3}, Muzhen He⁵, Zeyan Xu⁶,
Arlene Weir⁷, Hong-Yu Zhou⁸, Tianyu Zhang^{1 2 3}, Frederieke H van Duijnhoven⁹,
Luyi Han^{2 3}, Xiaomei Li¹⁰, Anna D'Angelo¹¹, Valentina Longo¹¹, Zaiyi Liu^{12 13},
Jonas Teuwen¹⁴, Marleen Kok^{15 16}, Regina Beets-Tan^{1 2}, Hugo M Horlings¹⁷, Tao Tan¹⁸,
Ritse Mann^{2 3}



- **mammografia e MRI** + info cliniche
- **deep-learning** model
- **3300 pazienti** in 16 anni da 4 dataset



2. modelli predittivi



3T DCE-MRI Radiomics Improves Predictive Models of Complete Response to Neoadjuvant Chemotherapy in Breast Cancer

Stefania Montemezzi¹, Giulio Benetti², Maria Vittoria Bisighin¹, Lucia Camera¹, Chiara Zerbato¹, Francesca Caumo³, Elena Fiorio⁴, Sara Zanelli⁴, Michele Zuffante⁵ and Carlo Cavedon^{2*}

¹ Radiology Unit, Azienda Ospedaliera Universitaria Integrata, Verona, Italy, ² Medical Physics Unit, Azienda Ospedaliera Universitaria Integrata, Verona, Italy, ³ Radiology Unit, Istituto Oncologico Veneto – IRCCS, Padova, Italy, ⁴ Pathology Unit, Azienda Ospedaliera Universitaria Integrata, Verona, Italy, ⁵ Nuclear Medicine Unit, Azienda Ospedaliera Universitaria Integrata, Verona, Italy

ORIGINAL RESEARCH
published: 20 April 2021
doi: 10.3389/fonc.2021.630780

- 20 complete responders (pCR) / 40 non-CR
- **radiomica + altri parametri quantitativi**
 - enhancement – type of curve (I-II-III)
 - ADC
 - PET-CT SUV_{max}
 - Ki-67 / ER / PgR / HER-2
 - type of lesion
 - grade / shape / int enhancement
- 5 radiomic features (geometric – 1st order – higher order)
- **classificazione con machine learning**
 - *logistic regression*
 - *support vector machines regression*
 - *random forest*

2. modelli predittivi

- *feature* di radiomica da sole danno prestazioni inferiori ai parametri istopatologici da soli
- l'inclusione nei modelli di *feature* di radiomica migliora la prestazione di tutti i modelli
- AUC > 0.9 con tutti i classificatori studiati
- => **l'analisi radiomica in MRI fornisce informazioni complementari (non ridondanti)**

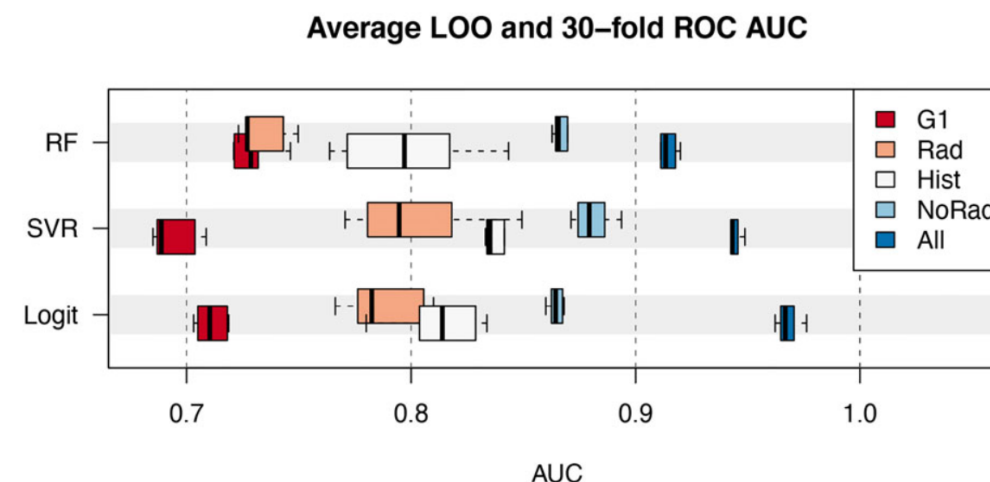


FIGURE 3 | Average AUC of the best 6 models as a function of the class of variables (G1, Rad, Hist, NoRad and All) and classifier (RF, random forest; SVR, Support Vector machines Regression; Logit, Logistic regression). Boxplots represent the median value, interquartile range and extremes.

La qualità dei dati

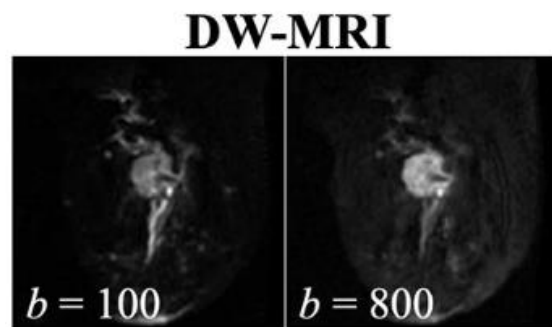
- **dati di training e di test / validazione**
 - responsabilità dello sviluppatore / fornitore
 - responsabilità del *deployer*
- **dati di input in fase di utilizzo**
 - responsabilità dell'utilizzatore
 - *molto importante se si impiegano dati quantitativi*



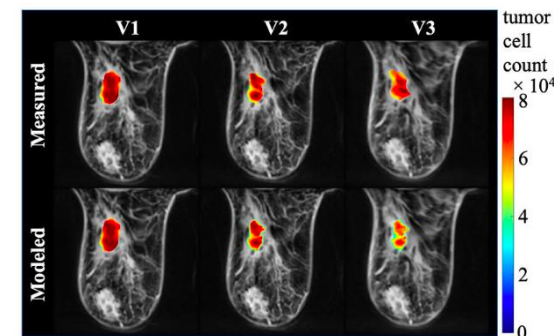
La qualità dei dati

- *standardizzazione*
 - dati omogenei in partenza
- *omogeneizzazione*
 - dati resi confrontabili (binning, resampling...)
- *raccolta e conservazione*
 - privacy
 - accessibilità
 - integrità
 - backup/conservazione

Ad es. in **modelli predittivi di risposta alla terapia** basati su dati quantitativi in MRI – valore di ADC *



$$N(\mathbf{x}, t) = \theta \left(\frac{ADC_w - ADC(\mathbf{x}, t)}{ADC_w - ADC_{min}} \right), \quad \rightarrow$$

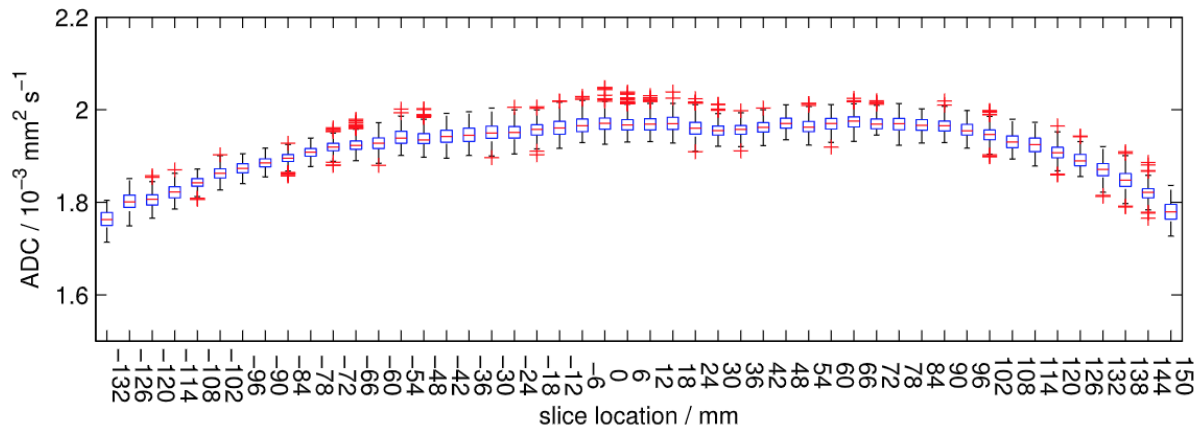


* Wu C, Jarrett AM, Zhou Z, Elshafeey N, et al. MRI-Based Digital Models Forecast Patient-Specific Treatment Responses to Neoadjuvant Chemotherapy in Triple-Negative Breast Cancer. **Cancer Res.** 2022 Sep 16;82(18):3394-3404. doi: 10.1158/0008-5472.CAN-22-1329. PMID: 35914239; PMCID: PMC9481712.

La qualità dei dati

- *standardizzazione*
 - dati omogenei in partenza
- *omogeneizzazione*
 - dati resi confrontabili (binning, resampling...)
- variabilità inter-scanner
- influenza di parametri quali ad es. posizione della slice **

- *raccolta e conservazione*
 - privacy
 - accessibilità
 - integrità
 - backup/conservazione

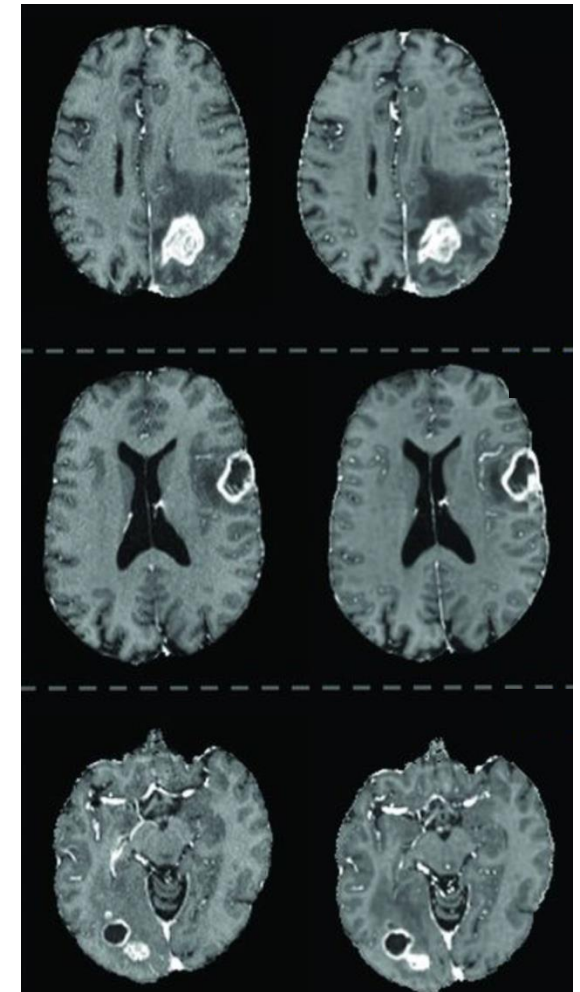
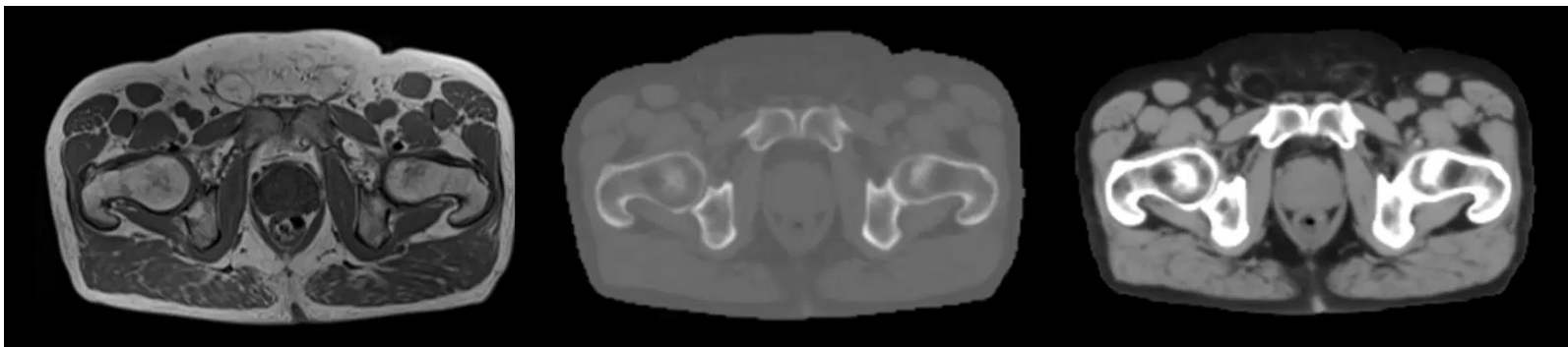


necessità di **standardizzazione** e **garanzia di qualità** a un **livello più elevato** rispetto a quanto necessario per applicazioni tradizionali

** JM Winfield, DJ Collins, AN Priest et al. "A framework for optimization of diffusion-weighted MRI protocols for large field-of-view abdominal-pelvic imaging in multicenter studies". Med. Phys. 43(1), pp. 95-110 (2016).

3. AI generativa

- algoritmi in grado di **generare contenuto informativo** (testo, immagini, suoni...)
- esempi:
 - **RM con contrasto virtuale** – permette di non somministrare mezzi di contrasto paramagnetici ad es. a base di gadolinio
 - **TC virtuale in radioterapia** – utilizzata per il calcolo della dose in radioterapia
 - **generazione di testo** (ausilio nella refertazione, workflow ecc.)

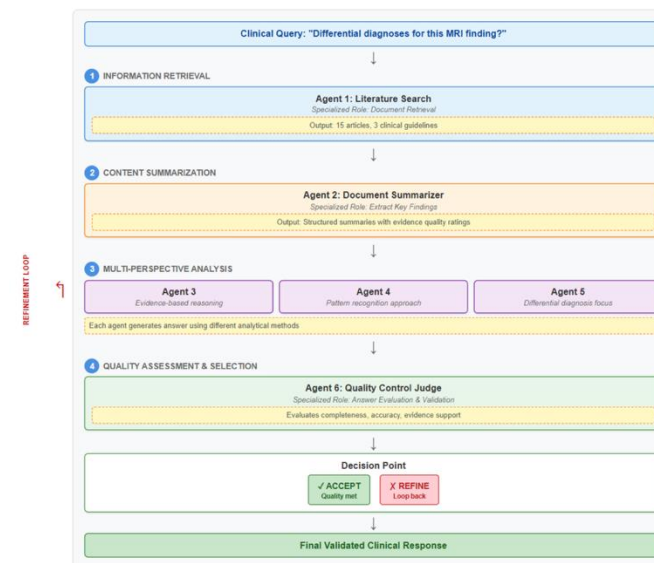


3. Al generativa: allucinazioni

- tendenza degli algoritmi generativi (in particolare LLM): **generare associazioni tra concetti correlati anche quando le prove disponibili sono limitate o ambigue**
- pericolo nelle applicazioni cliniche: tendenza a **produrre risposte apparentemente plausibili piuttosto che riconoscere adeguatamente l'incertezza** di fronte a informazioni ambigue
- => servono un **programma di garanzia di sicurezza e qualità** dedicato e **strategie di mitigazione**
 - della tendenza a **produrre risultati non correlati ai vincoli** dello specifico problema analizzato
 - della tendenza a **produrre un risultato in ogni caso**

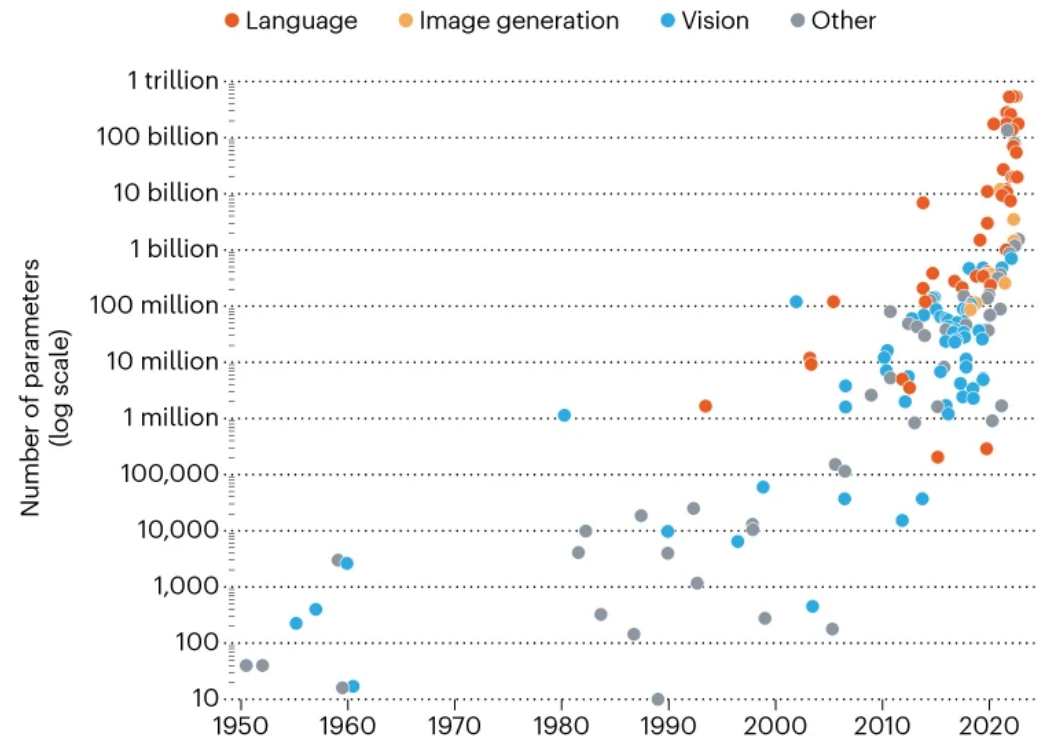
3. AI generativa: allucinazioni

- strategie basate su **correzione del modello dopo identificazione delle allucinazioni**
- vantaggio: **correzione guidata da agente umano**
- svantaggi: **non previene le cause** ma corregge a posteriori
- **retrieval-augmented generation (RAG)**
- basa le risposte su fonti esplicite e verificabili, riducendo il rischio di allucinazioni
- **noscendo le allucinazioni ma in modo preventivo**
- schemi che suddividono i compiti di un LLM in più stadi: **multi-agent frameworks** (ad es. Agentic AI)
- **checkpoint di validazione esterna** tra le fasi di elaborazione



4. Algoritmi evolutivi e *foundation models*

- "models trained on broad data that can be adapted to a wide range of downstream tasks"*
- modelli **generalizzabili** (general-purpose AI - **GPAI**)
- **architettura comune** ad elevate prestazioni che viene adattata (**fine tuning**) a problemi specifici modificando solo una piccola parte del modello (vs. narrow-AI)
- **altissimo numero di parametri**
- richiedono training su database di enormi dimensioni => **costi elevati** (1-10 M\$)



Nature, news section, 14 March 2023

Adapted from: J. Sevilla et al. Preprint at <https://arxiv.org/abs/2202.05924> (2022)

* Stanford Institute for Human-Centered Artificial Intelligence, 'Reflections on Foundation Models'
<<https://hai.stanford.edu/news/reflections-foundation-models>>, accessed 8 June 2026.

4. Algoritmi evolutivi e *foundation models*

Article

A fully open AI foundation model applied to chest radiography

<https://doi.org/10.1038/s41586-025-09079-8>

DongAo Ma¹, Jiaxuan Pang¹, Michael B. Gotway² & Jianming Liang^{3,✉}

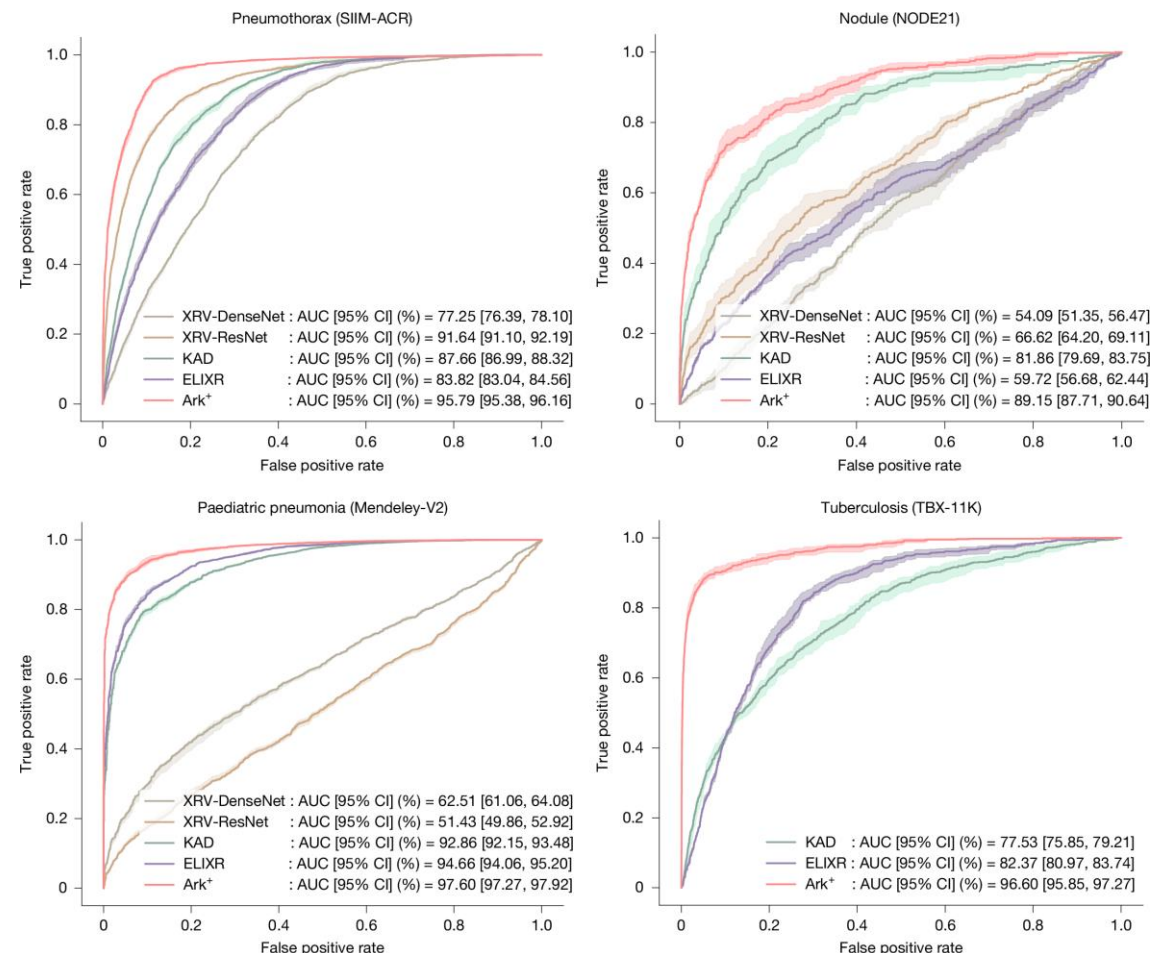
Received: 6 July 2024

Accepted: 29 April 2025

Published online: 11 June 2025

488 | Nature | Vol 643 | 10 July 2025

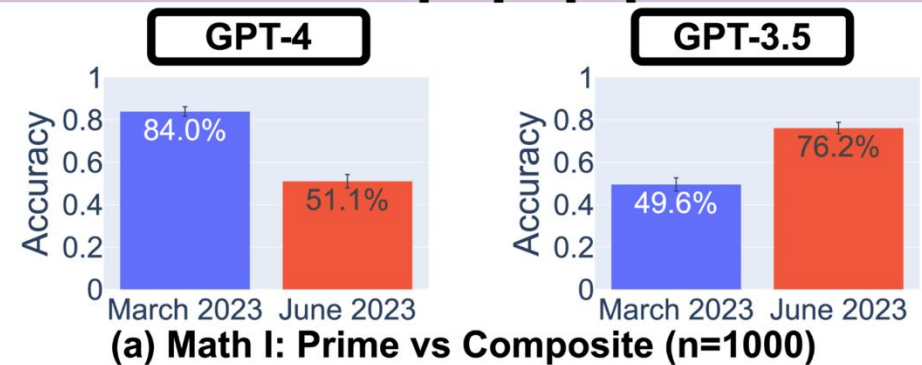
- esempio di *foundation model* in radiologia
- algoritmo evolutivo – *continuous learning*
- mantenimento delle prestazioni nel tempo: *stress test* con periodicità programmata



4. Algoritmi evolutivi e *foundation models*

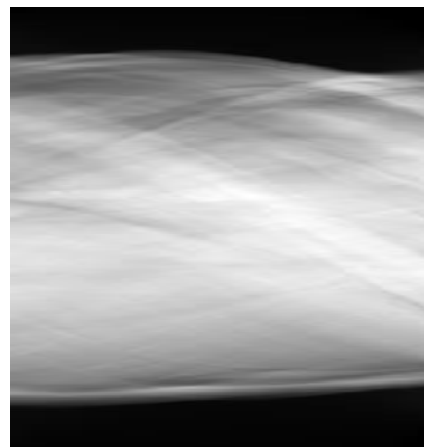
- processo **non completamente tracciabile** a causa di
 - complessità
 - evoluzione continua (life-long learning)
- possibile **degrado di prestazioni** imprevedibile a causa dell'adattamento continuo ai dati
- necessario **garantire robustezza degli algoritmi** – programma di *stress test* periodici

Is 17077 a prime number? Think step by step and then answer [Yes] or [No].



5. Lettura diretta di *raw data*

I "dati grezzi" a monte dell'immagine contengono maggiore informazione dell'immagine ricostruita



- ***raw data* impiegati da algoritmi di deep learning:** potenzialità di sfruttare tutta l'informazione presente
- criticità: ridotta possibilità di interpretazione / **validazione umana**

5. Lettura diretta di *raw data*

In the Era of Deep Learning, Why Reconstruct an Image at All?

Caroline Chung, MD, MSc, CIP, Jayashree Kalpathy-Cramer, PhD, Michael V. Knopp, MD, David A. Jaffray, PhD



Journal of the American College of Radiology

Volume 18, Issue 1, Part B, January 2021, Pages 170-173



Phys. Med. Biol. 69 (2024) 075015

<https://doi.org/10.1088/1361-6560/ad1e7c>

Physics in Medicine & Biology

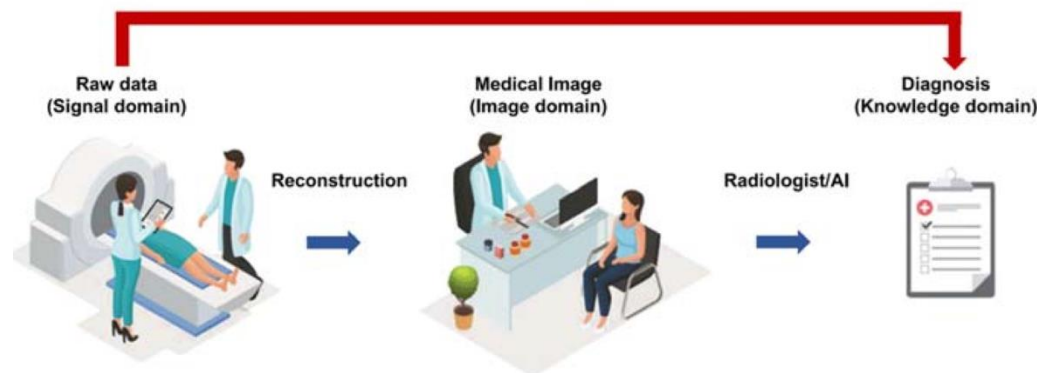


PAPER

Breaking boundaries in radiology: redefining AI diagnostics via raw data ahead of reconstruction

Bingxi He^{1,2,3,9}, Caixia Sun^{1,2,3,9}, Hailin Li^{1,2,3,9}, Yongbo Wang^{4,9}, Yunlang She⁵, Mengmeng Zhao⁵, Mengjie Fang^{1,2,3}, Yongbei Zhu^{1,2,3}, Kun Wang³, Zhenyu Liu³, Ziqi Wei³, Wei Mu^{1,2,3}, Shuo Wang^{1,2,3}, Zhenchao Tang^{1,2,3}, Jingwei Wei³, Lizhi Shao³, Lixia Tong⁶, Feng Huang⁶, Mingze Tang⁷, Yu Guo⁸, Huimao Zhang⁸, Di Dong^{3,*}, Chang Chen^{5,*}, Jianhua Ma^{4,*} and Jie Tian^{1,2,3,*}

Raw data Analysis



5. Lettura diretta di *raw data*

www.nature.com/scientificreports

**SCIENTIFIC
REPORTS**
nature research

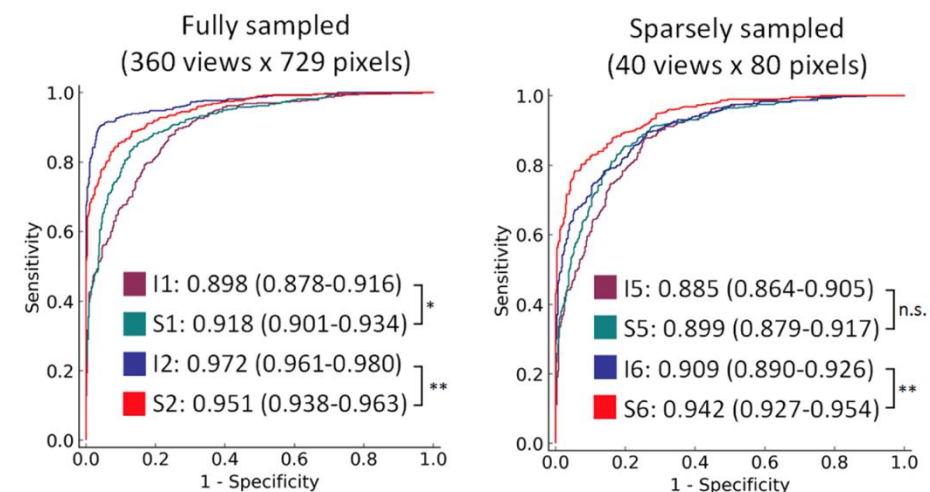
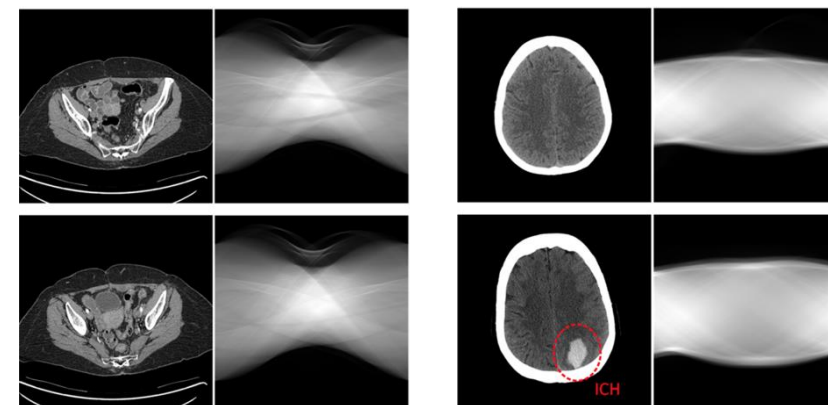
OPEN

Machine Friendly Machine Learning: Interpretation of Computed Tomography Without Image Reconstruction

Hyunkwang Lee^{1,2}, Chao Huang¹, Sehyo Yune¹, Shahein H. Tajmir¹, Myeongchan Kim¹ & Synho Do^{1*}

SCIENTIFIC REPORTS | (2019) 9:15540 | <https://doi.org/10.1038/s41598-019-51779-5>

- training su immagini ricostruite e su sinogrammi CT
 - *riconoscimento dettagli anatomici*
 - *identificazione di anomalie*
- prestazioni superiori su immagine se *fully sampled*
- prestazioni superiori su sinogrammi se *sparsely sampled*



Conclusioni

- la capacità degli algoritmi di IA di estrarre informazioni anche "nascoste" richiede di **innalzare e controllare il livello della qualità dei dati in ingresso**
- l'uso di algoritmi evolutivi (ma anche gli aggiornamenti successivi di algoritmi statici) richiede di garantire la **robustezza delle prestazioni**
- lo sviluppo di algoritmi ad alte prestazioni è sempre più riservato a pochi player con grandi risorse ma la **gestione di sicurezza e qualità è a tutti i livelli**, compreso l'utilizzatore finale
- in AOUI Verona **team multidisciplinare** dedicato all'implementazione dell'intelligenza artificiale nei processi aziendali (*team IA*)

