PROF. FEDERICO CABITZA

Dipartimento Informatica, Sistemistica e Comunicazione,
Università degli Studi Milano Bicocca

POTENZIALITÀ E RISCHI DELLA APPLICAZIONE DELLA INTELLIGENZA ARTIFICIALE IN MEDICINA

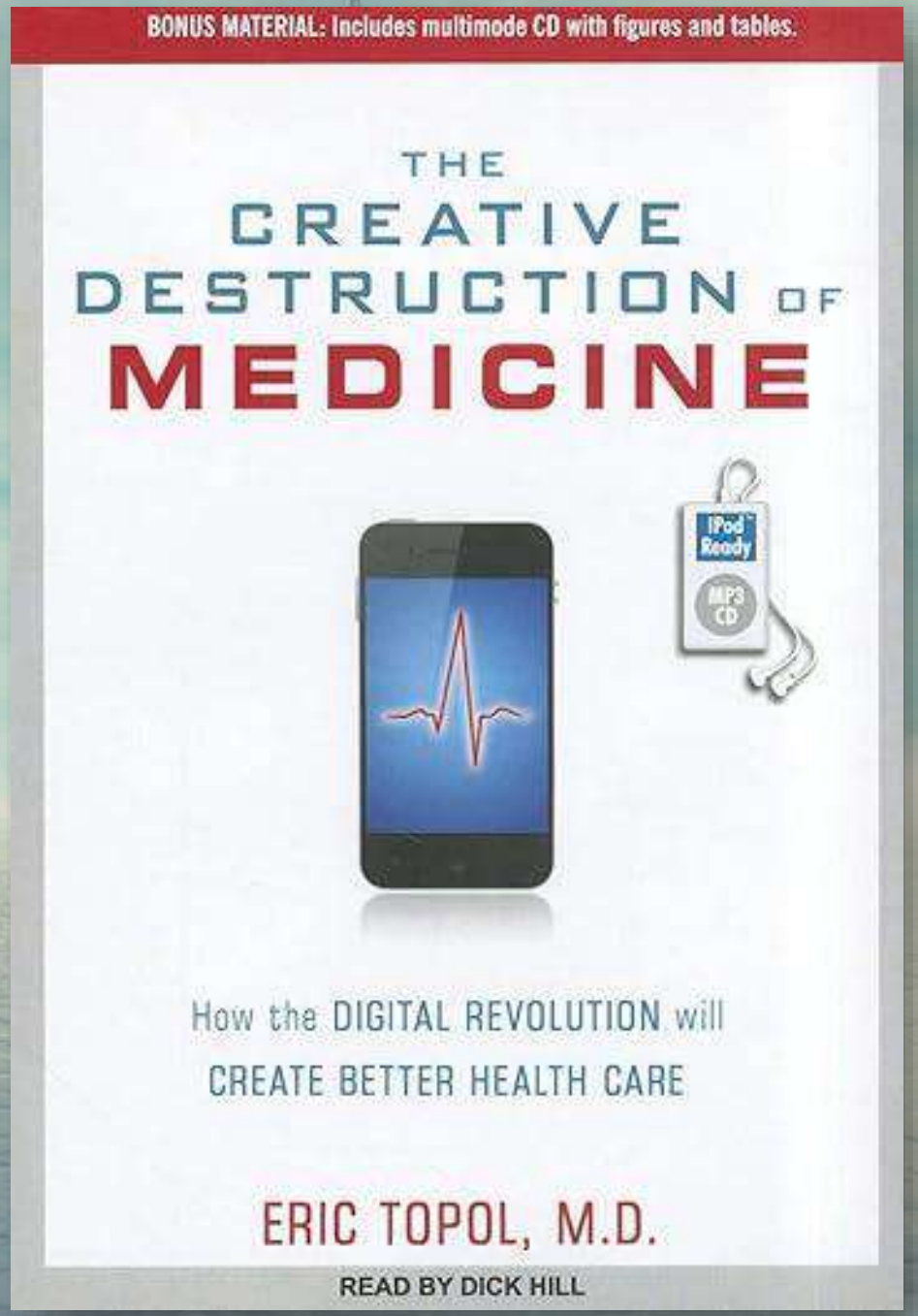
PROF. FEDERICO CABITZA

UNIVERSITY OF MILANO-BICOCCA

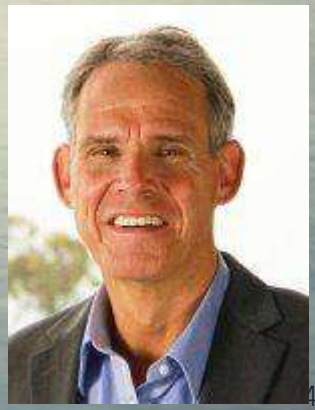
VIALE SARCA 336, MILAN, ITALY

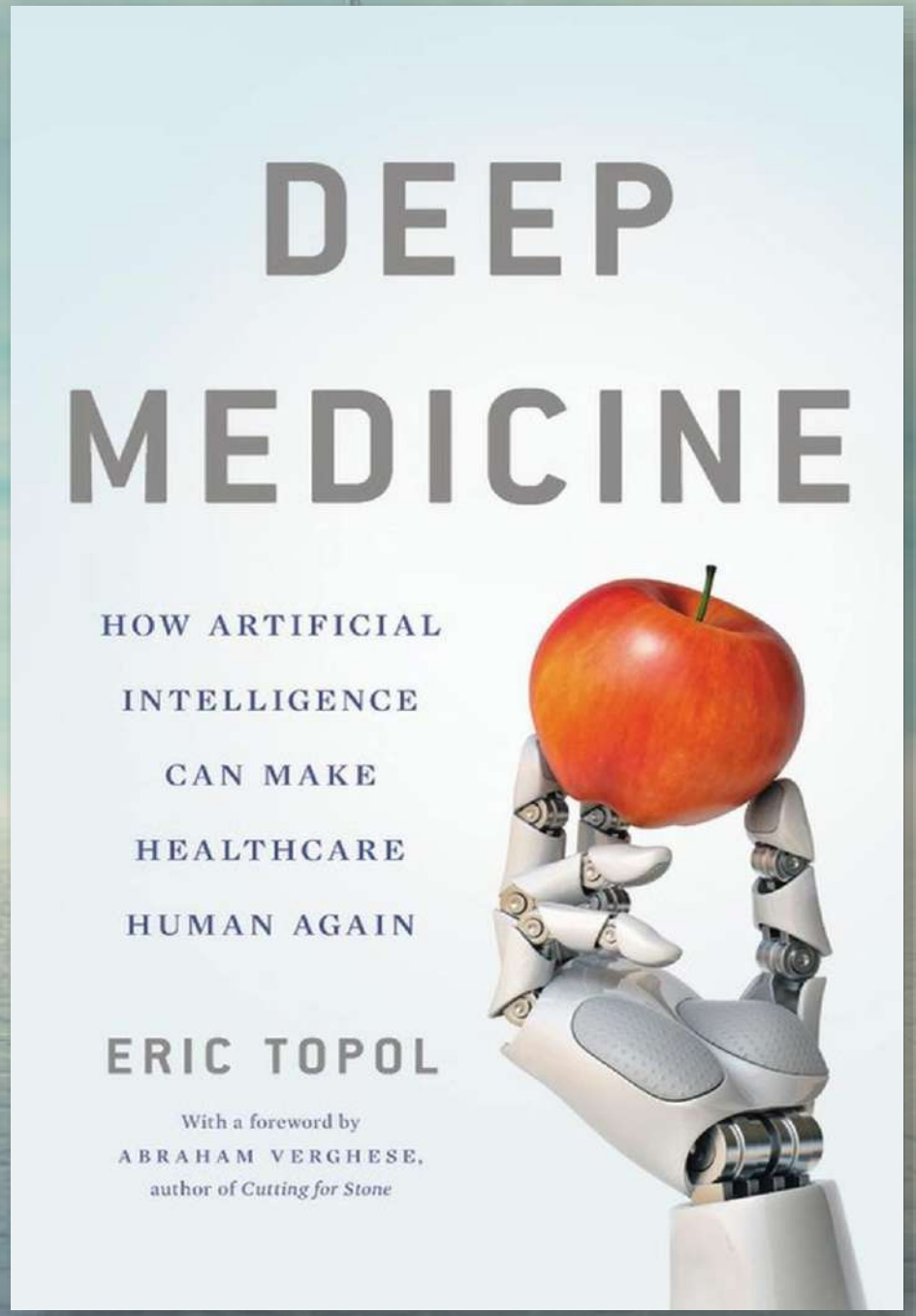
IRCSS HOSPITAL GALEAZZI-SANT'AMBROGIO, MILAN, ITALY


2012-2022
2023-2033

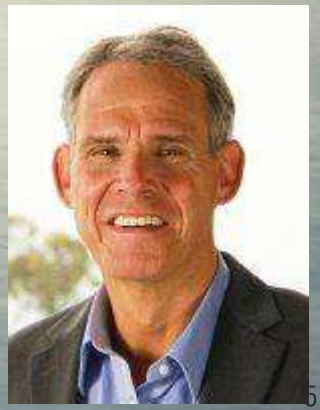

BONUS MATERIAL: Includes multimode CD with figures and tables.
THE
CREATIVE
DESTRUCTION OF
MEDICINE
How the DIGITAL REVOLUTION will
CREATE BETTER HEALTH CARE
ERIC TOPOL, M.D.
READ BY DICK HILL

Marzo 2012




DEEP
MEDICINE
HOW ARTIFICIAL
INTELLIGENCE
CAN MAKE
HEALTHCARE
HUMAN AGAIN
ERIC TOPOL
With a foreword by
ABRAHAM VERGHESE,
author of Cutting for Stone

Luglio 2019



VIEWPOINT

Unintended Consequences of Machine Learning in Medicine

Federico Cabitza, PhD
Department of Informatics, University of Milano-Bicocca, Milan, Italy; and IRCCS Istituto Ortopedico Galeazzi, Milan, Italy.

Raffaele Rasoini, MD
Centro Studi Medicina Avanzata, Florence, Italy.

Gian Franco Gensini, MD
Centro Studi Medicina Avanzata, Florence, Italy.

Over the past decade, machine learning techniques have made substantial advances in many domains. In health care, global interest in the potential of machine learning has increased; for example, a deep learning algorithm has shown high accuracy in detecting diabetic retinopathy.[1] There have been suggestions that machine learning will drive changes in health care within a few years, specifically in medical disciplines that require more accurate prognostic models (eg, oncology) and those based on pattern recognition (eg, radiology and pathology).

However, comparative studies on the effectiveness of machine learning–based decision support systems (ML-DSS) in medicine are lacking, especially regarding the effects on health outcomes. Moreover, the introduction of new technologies in health care has not always been straightforward or without unintended and adverse effects.[2] In this Viewpoint we consider the potential the expense of other elements that are more difficult or impossible to easily describe. Relying on ML-DSS requires considering digital data as reliable and complete representations of the phenomena that these data are supposed to render in a discrete and trustv
This may be a problem when the clinical c
represented, particularly if physicians lose a
the existence of clinical elements that are not included in the clinical record.

Such lack of information may lead to partial or misleading interpretations of ML-DSS diagnostics and therapeutic or prognostic outputs. It also could lead to reduced interest in and decreased ability to perform holistic evaluations of patients, with loss of valuable and irreducible aspects of the human experience such as psychological, relational, social, and organizational issues. These factors may not be incorporated into any ML-DSS

Agosto 2017

VIEWPOINT

Unintended Consequences of Machine Learning in Medicine

Federico Cabitza, PhD
Department of Informatics, University of Milano-Bicocca, Milan, Italy; and IRCCS Istituto Ortopedico Galeazzi, Milan, Italy.

Raffaele Rasoini, MD
Centro Studi Medicina Avanzata, Florence, Italy.

Gian Franco Gensini, MD
Centro Studi Medicina Avanzata, Florence, Italy.

Over the past decade, machine learning techniques have made substantial advances in many domains. In health care, global interest in the potential of machine learning has increased; for example, a deep learning algorithm has shown high accuracy in detecting diabetic retinopathy.[1] There have been suggestions that machine learning will drive changes in health care within a few years, specifically in medical disciplines that require more accurate prognostic models (eg, oncology) and those based on pattern recognition (eg, radiology and pathology).

However, comparative studies on the effectiveness of machine learning–based decision support systems (ML-DSS) in medicine are lacking, especially regarding the effects on health outcomes. Moreover, the introduction of new technologies in health care has not always been straightforward or without unintended and adverse effects.[2] In this Viewpoint we consider the po-

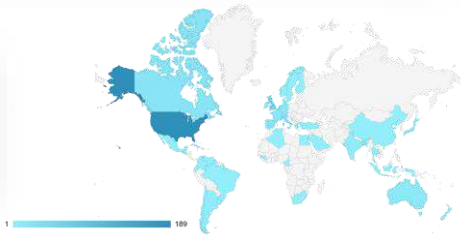
the expense of other elements that are more difficult or impossible to easily describe. Relying on ML-DSS requires considering digital data as reliable and complete representations of the phenomena that these data are supposed to render in a discrete and trustv
This may be a problem when the clinical c
represented, particularly if physicians lose a
the existence of clinical elements that are not included in the clinical record.

Such lack of information may lead to partial or misleading interpretations of ML-DSS diagnostics and therapeutic or prognostic outputs. It also could lead to reduced interest in and decreased ability to perform holistic evaluations of patients, with loss of valuable and irreducible aspects of the human experience such as psychological, relational, social, and organizational issues. These factors may not be incorporated into any ML-DSS

Agosto 2017

In the top 5% of all research outputs scored by Altmetric

High Attention Score compared to outputs of the same age (99th percentile)



VIEWPOINT

Unintended Consequences of Machine Learning in Medicine

EROSIONE DELLE COMPETENZE

(deskilling)

VIEWPOINT

Unintended Consequences of Machine Learning in Medicine

SVALUTAZIONE DEL CONTESTO

(demise of context)

VIEWPOINT

Unintended Consequences of Machine Learning in Medicine

SOTTOSTIMA INCERTEZZA E AMBIGUITÀ

Health Informatics Journal
Volume 25, Issue 3, September 2019, Pages 475-480
© The Author(s) 2019, Article Reuse Guidelines
https://doi.org/10.1177/1460458219824705

SAGE journals

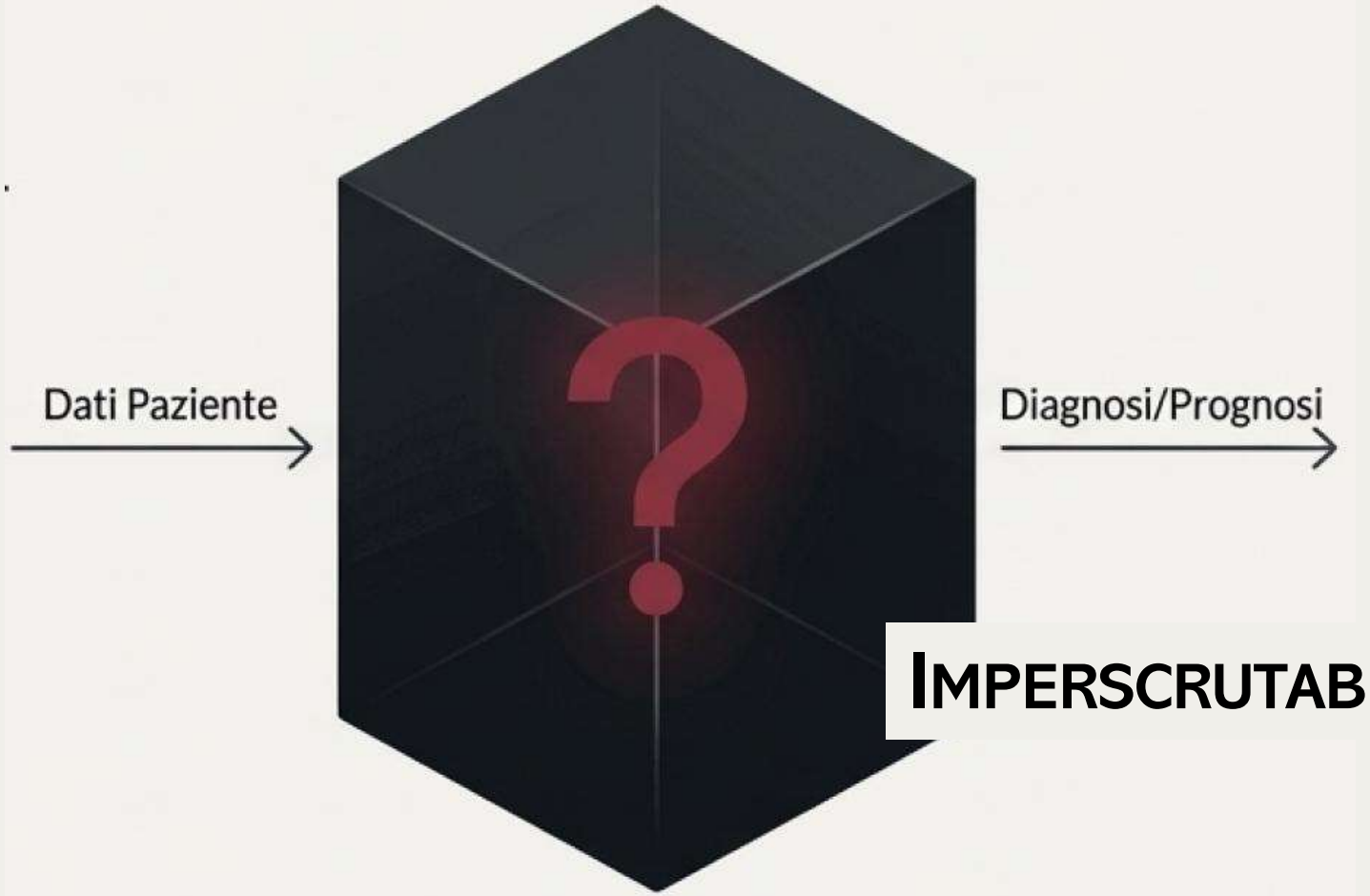
Special Issue Article

The elephant in the record: On the multiplicity of data recording work

Federico Cabitza [1], Angela Locoro [2], Camilla Alderighi, Raffaele Rasoini [3], Domenico Compagnone, and Pedro Berjano [4]

VIEWPOINT

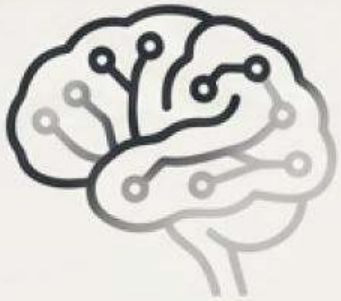
Unintended Consequences of Machine Learning in Medicine



IMPERSCRUTABILITÀ

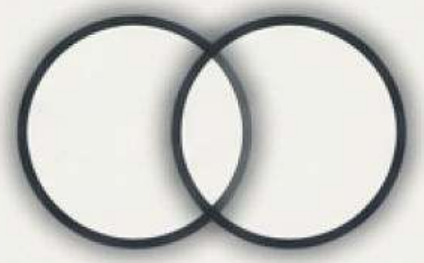
VIEWPOINT

Unintended Consequences of Machine Learning in Medicine



EROSIONE DELLE COMPETENZE

SVALUTAZIONE DEL CONTESTO

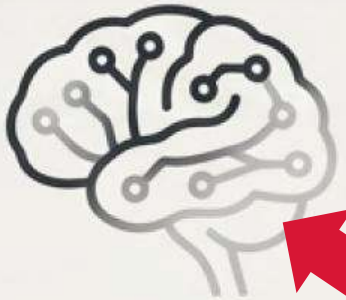


SOTTOSTIMA INCERTEZZA E AMBIGUITÀ

IMPERSCRUTABILITÀ

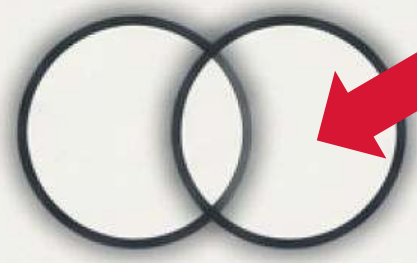
VIEWPOINT

Unintended Consequences of Machine Learning in Medicine

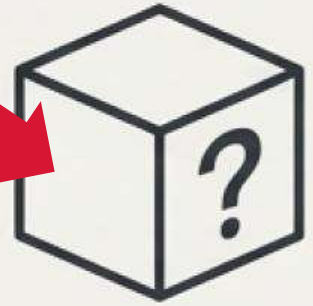


EROSIONE DELLE COMPETENZ

SVALUTAZIONE DEL CONTESTO



SOTTOSTIMA INCERTEZZA E AMBIGUITÀ



IMPERSCRUTABILITÀ

VIEWPOINT

Unintended Consequences of Machine Learning in Medicine

*"La qualità di qualsiasi ML-DSS... non dovrebbe essere fondata solo su metriche di performance, ma dovrebbe **piuttosto** essere soggetta alla **prova di miglioramenti clinicamente importanti negli esiti** rilevanti, rispetto alla cura abituale, insieme alla soddisfazione di pazienti e medici."*

Federico Cabitza et al, JAMA, 2017

4 giorni fa…

International Journal of Medical Informatics
Volume 207, 1 March 2026, 106220

Review article

Effectiveness of computerized decision support systems linked to electronic health records: An updated systematic review with *meta*-analysis

Annalisa Biffi [a], Greta Castellini [b], Gabriele del Castillo [c], Francesca De Nard [c], Camilla Vismara [a], Federico Cabitza [d e], Giovanni Corrao [a], Silvia Gianola [b]

AJPH
A PUBLICATION OF THE AMERICAN PUBLIC HEALTH ASSOCIATION

Home » American Journal of Public Health (AJPH) » December 2014

Effectiveness of Computerized Decision Support Systems Linked to Electronic Health Records: A Systematic Review and Meta-Analysis

Lorenzo Moja MD, MSc, PhD, Koren H. Kwag BSc, MSc, Theodore Lytras MD, MPH, Lorenzo Bertizzolo MD, Linn Brandt MD, Valentina Pecoraro BSc, Giulio Rigon MD, MSc, Alberto Vaona MD, MSc, Francesca Ruggiero BA, MA, Massimo Mangia, Alfonso Iorio MD, PhD, Ilkka Kunnamo MD, PhD, and Stefanos Bonovas MD, MSc, PhD

2013-2023

ELSEVIER

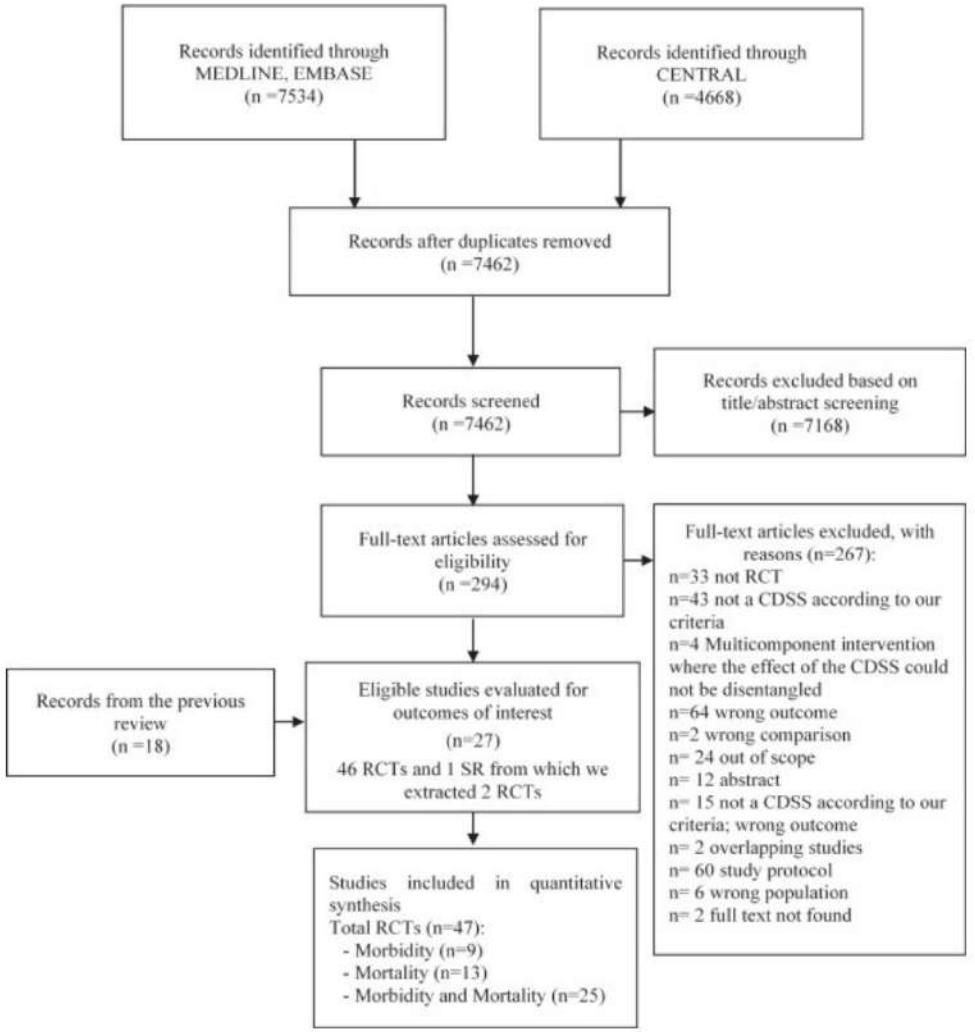
International Journal of Medical Informatics

Volume 207, 1 March 2026, 106220

Review article

Effectiveness of computerized decision support systems linked to electronic health records: An updated systematic review with *meta*-analysis

Annalisa Biffi [a] ✉, Greta Castellini [b] ✉, Gabriele del Castillo [c] ✉, Francesca De Nard [c] ✉, Camilla Vismara [a] ✉, Federico Cabitza [d e] ✉, Giovanni Corrao [a] ✉, Silvia Gianola [b] ✉

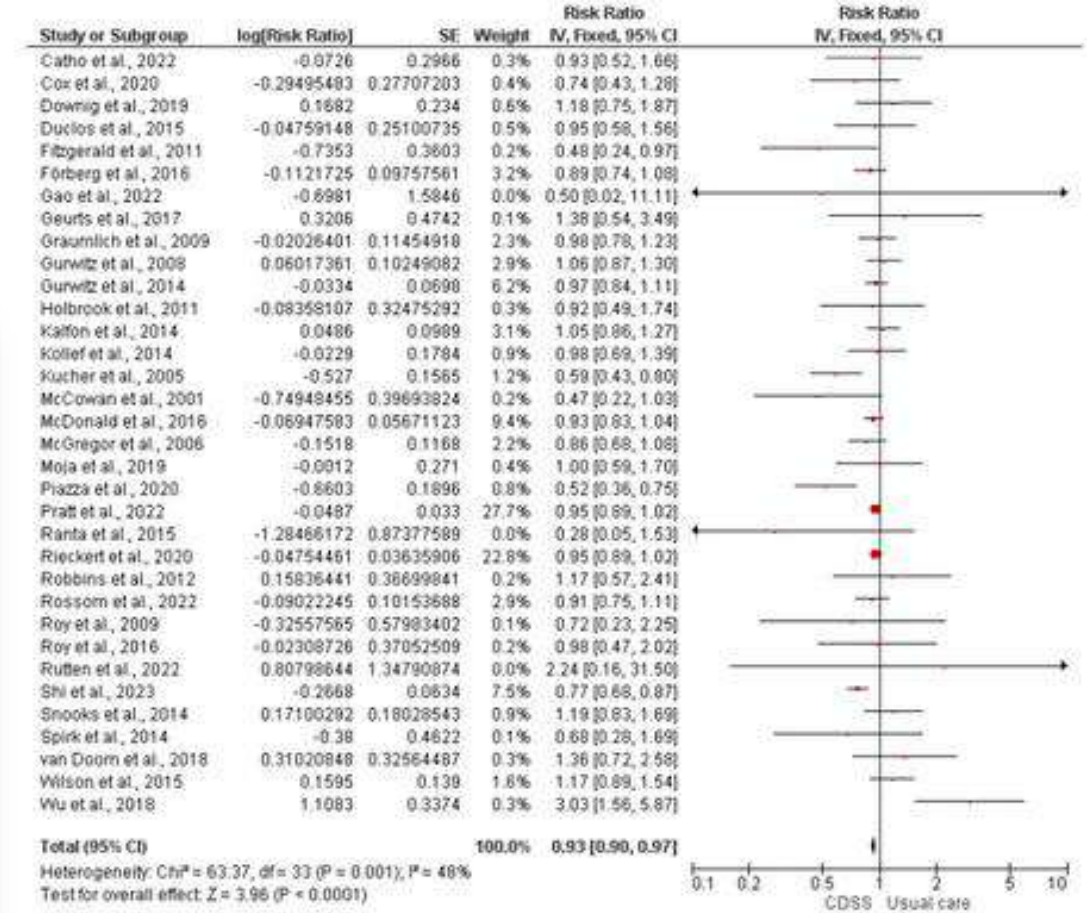


International Journal of Medical Informatics
Volume 207, 1 March 2026, 106220

Review article

Effectiveness of computerized decision support systems linked to electronic health records: An updated systematic review with *meta*-analysis

Annalisa Biffi [a], Greta Castellini [b], Gabriele del Castillo [c], Francesca De Nard [c], Camilla Vismara [a], Federico Cabitza [d e], Giovanni Corrao [a], Silvia Gianola [b]



554

555 **Figure 3.** Forest plots from individual studies and meta-analysis for morbidity, any

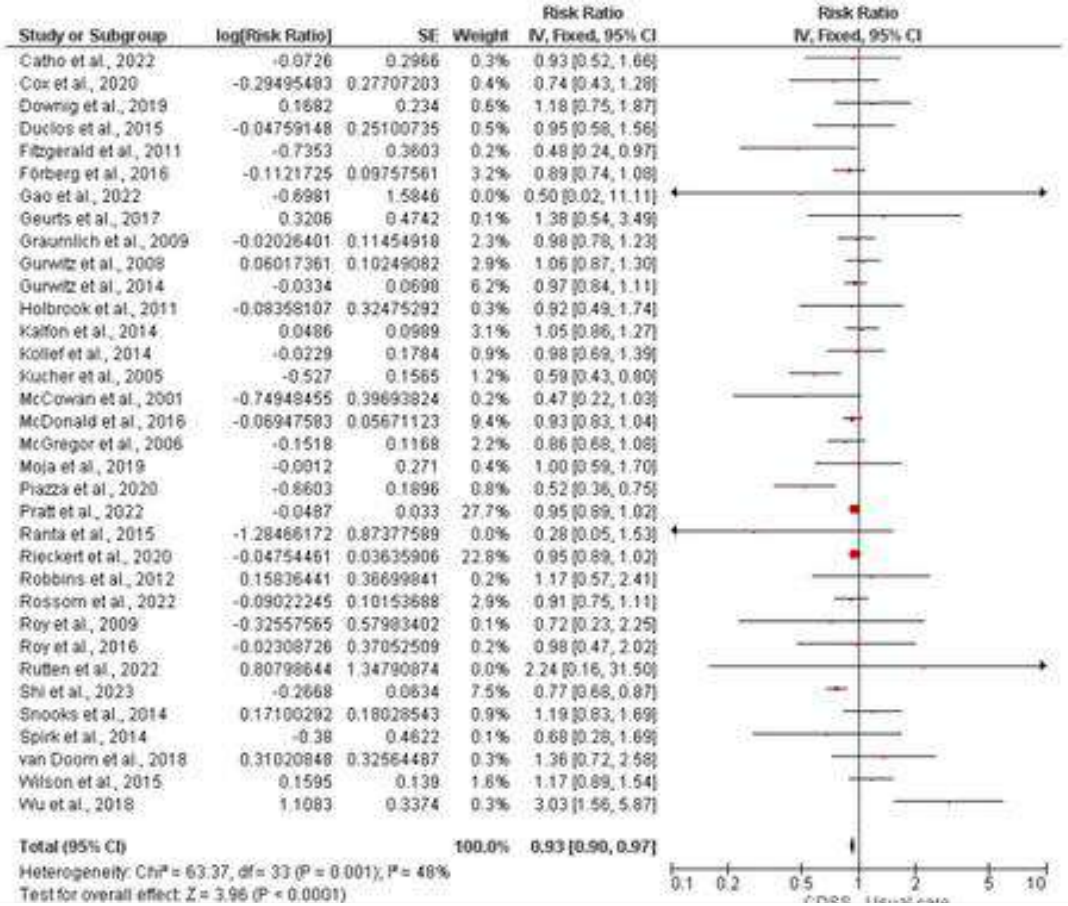
mortality

morbidity

International Journal of Medical Informatics
Volume 207, 1 March 2026, 106220

Review article

Effectiveness of computerized decision support systems linked to electronic health records: An updated systematic review with *meta*-analysis



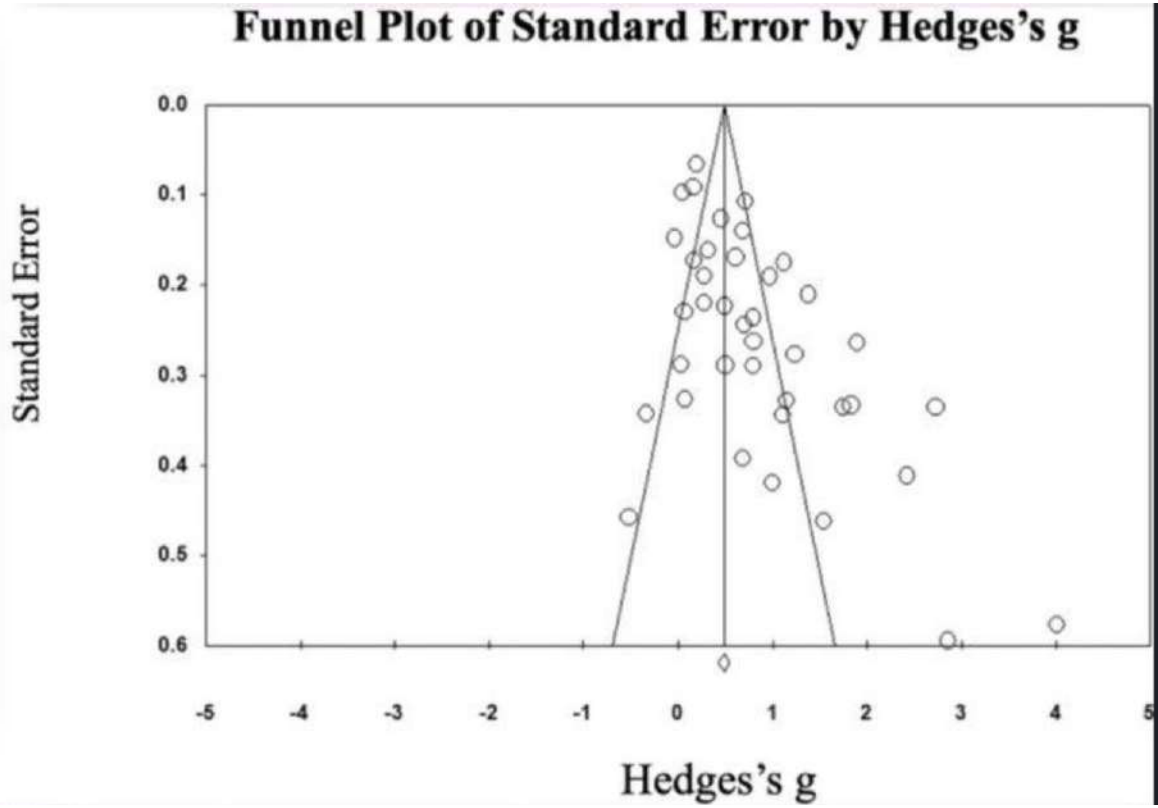
We included 47 RCTs, incorporating data from 29 new RCTs in this update. Compared with controls, CDSS use may result in little to no reduction in mortality (38 trials,127,623 patients; fixed-effects model risk ratio [RR] = 0.98; 95% confidence interval[CI] 0.93 to 1.02; I2 = 0%; moderate certainty). The meta-analysis on morbidity reached nominal statistical significance: CDSS use may have trivial or small benefits with respect to morbidity (34 RCTs; 133,504 patients; fixed-effects model RR = 0.92,95% CI 0.90–0.97; random-effects model RR = 0.93, 95% CI 0.87–0.99; I2 = 48%; high certainty). **Our meta-analysis did not highlight substantial effects on mortality while tiny reductions in morbidity are possible**. In specific therapeutic areas, such as cardiovascular, a small effect may be present.

International Journal of Medical Informatics
Volume 207, 1 March 2026, 106220

Review article

Effectiveness of computerized decision support systems linked to electronic health records: An updated systematic review with *meta*-analysis

Annalisa Biffi [a], Greta Castellini [b], Gabriele del Castillo [c], Francesca De Nard [c], Camilla Vismara [a], Federico Cabitza [d e], Giovanni Corrao [a], Silvia Gianola [b]



… e nei prossimi 7 anni?

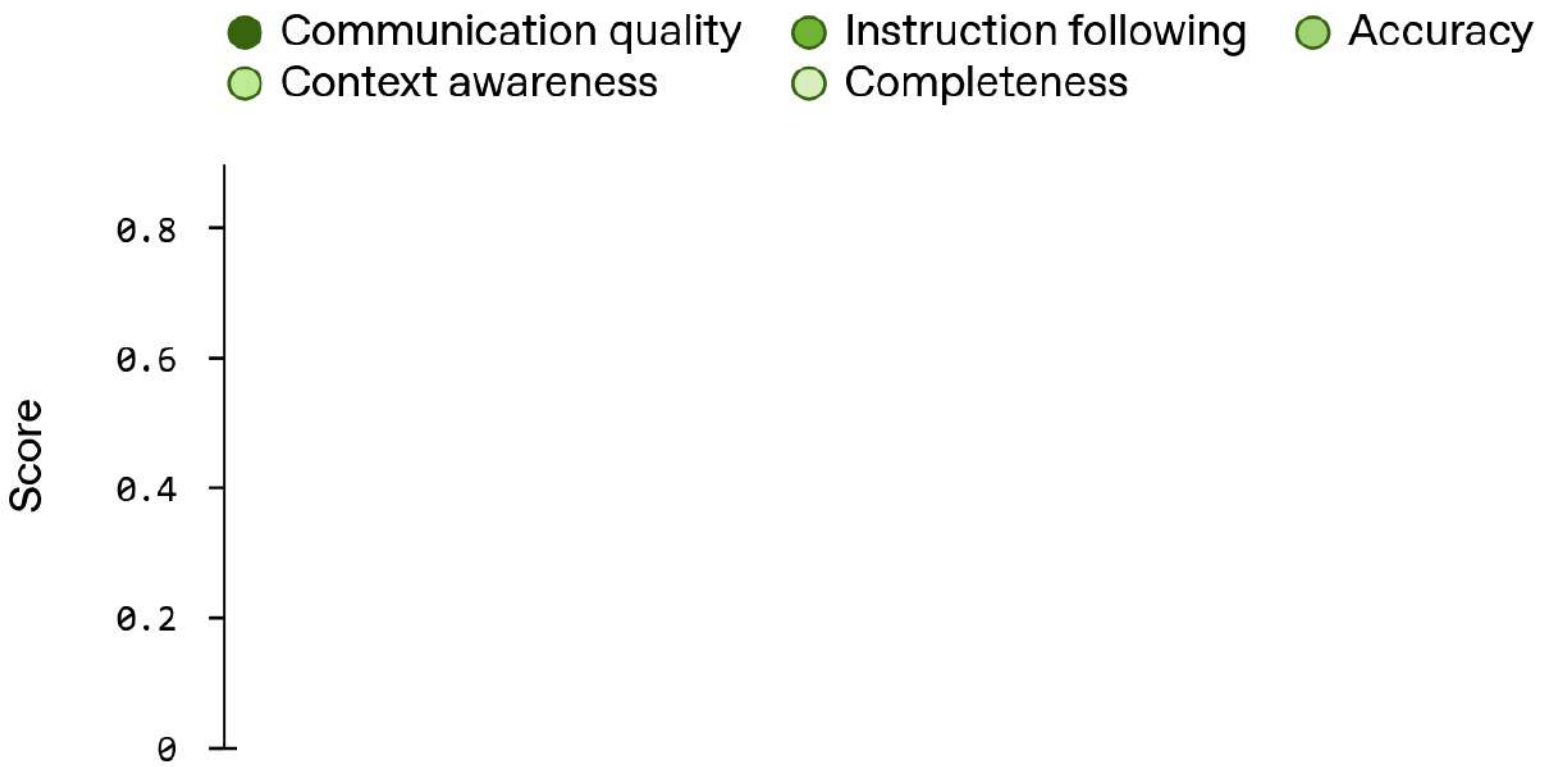
May 12, 2025 Publication

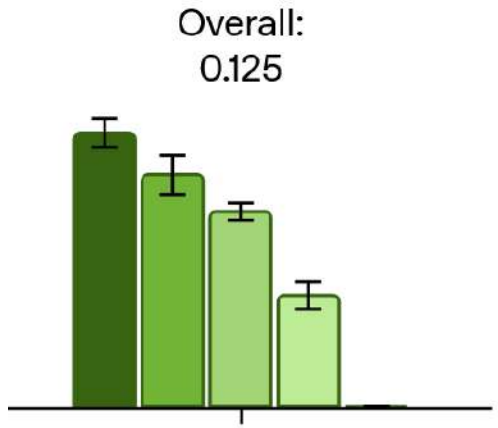
Introducing HealthBench

An evaluation for AI systems and human health.

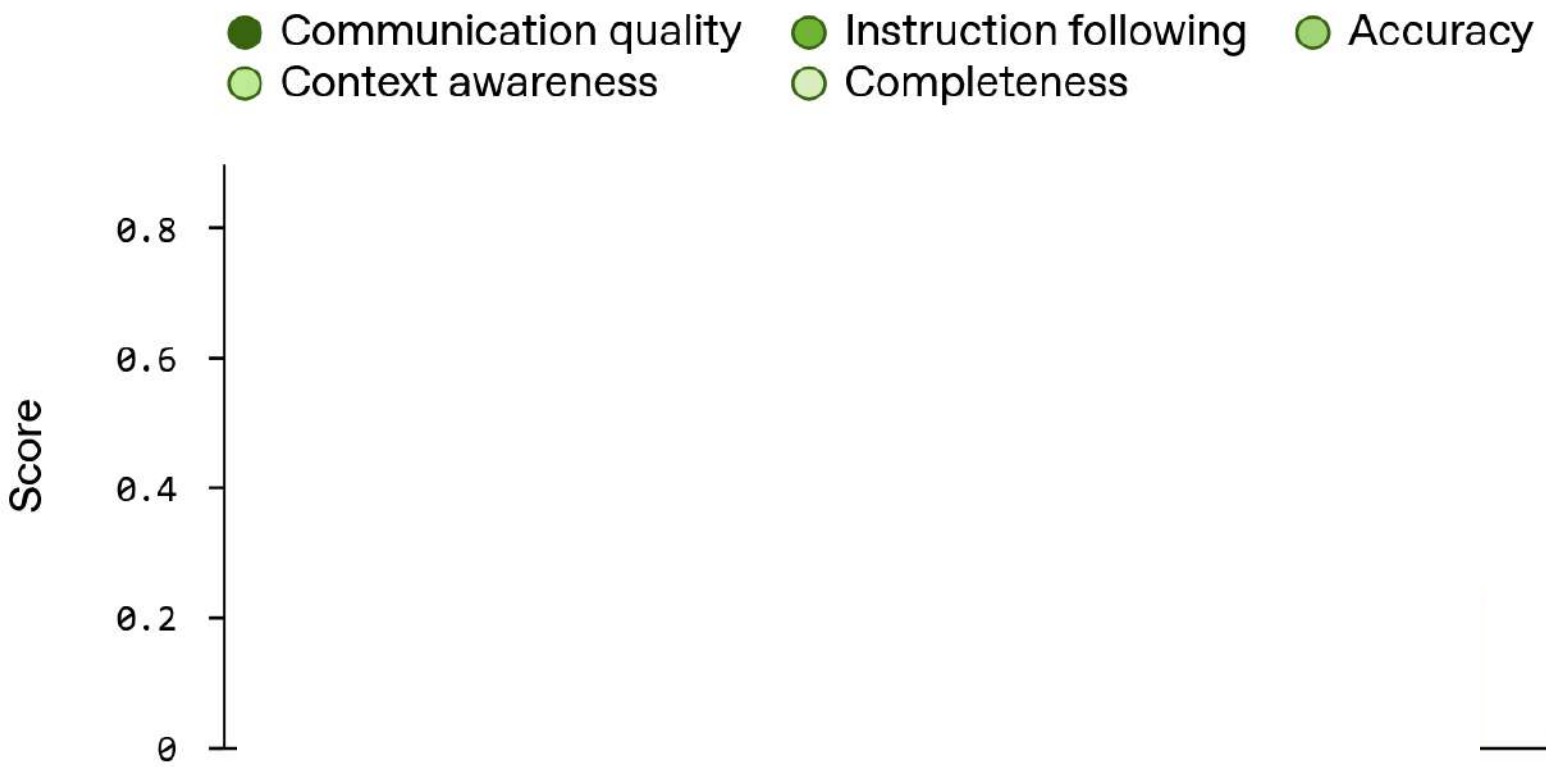
Valutazione risposte scritte dei medici

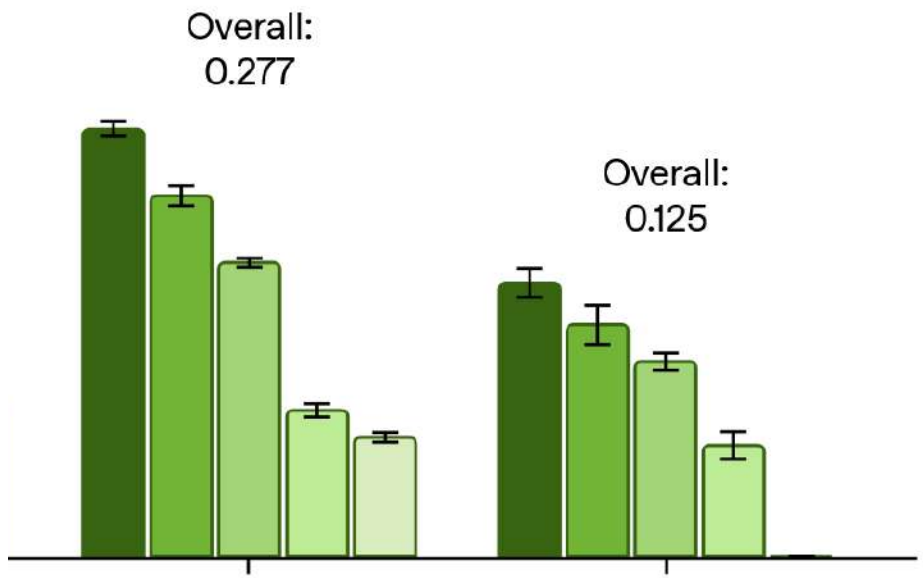
Valutazione risposte scritte dei medici

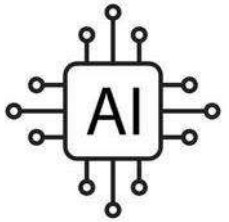




Valutazione risposte scritte dei medici

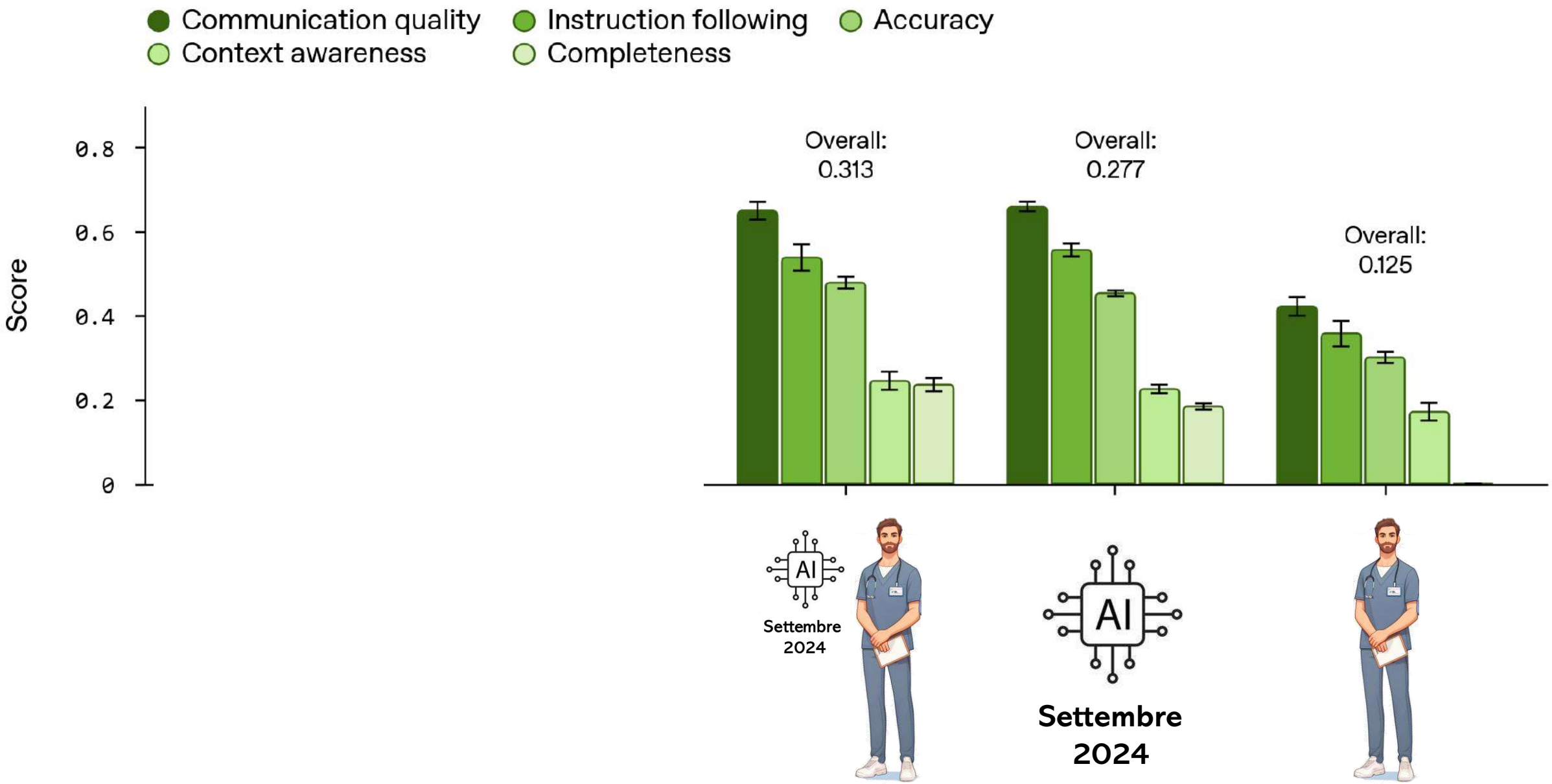




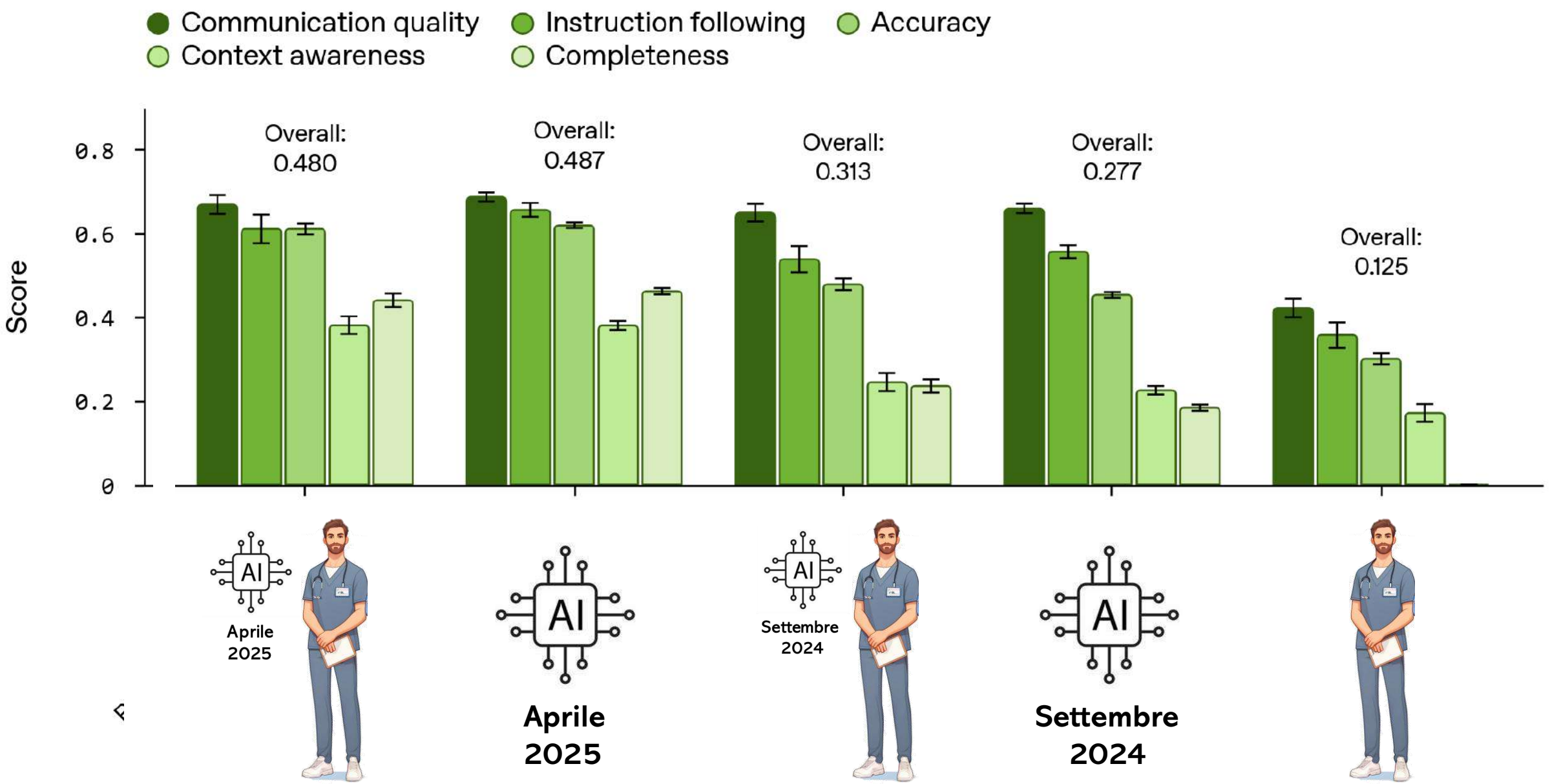


Settembre 2024

Valutazione risposte scritte dei medici



Valutazione risposte scritte dei medici



VIEWPOINT

Unintended Consequences of Machine Learning in Medicine

EROSIONE DELLE COMPETENZE

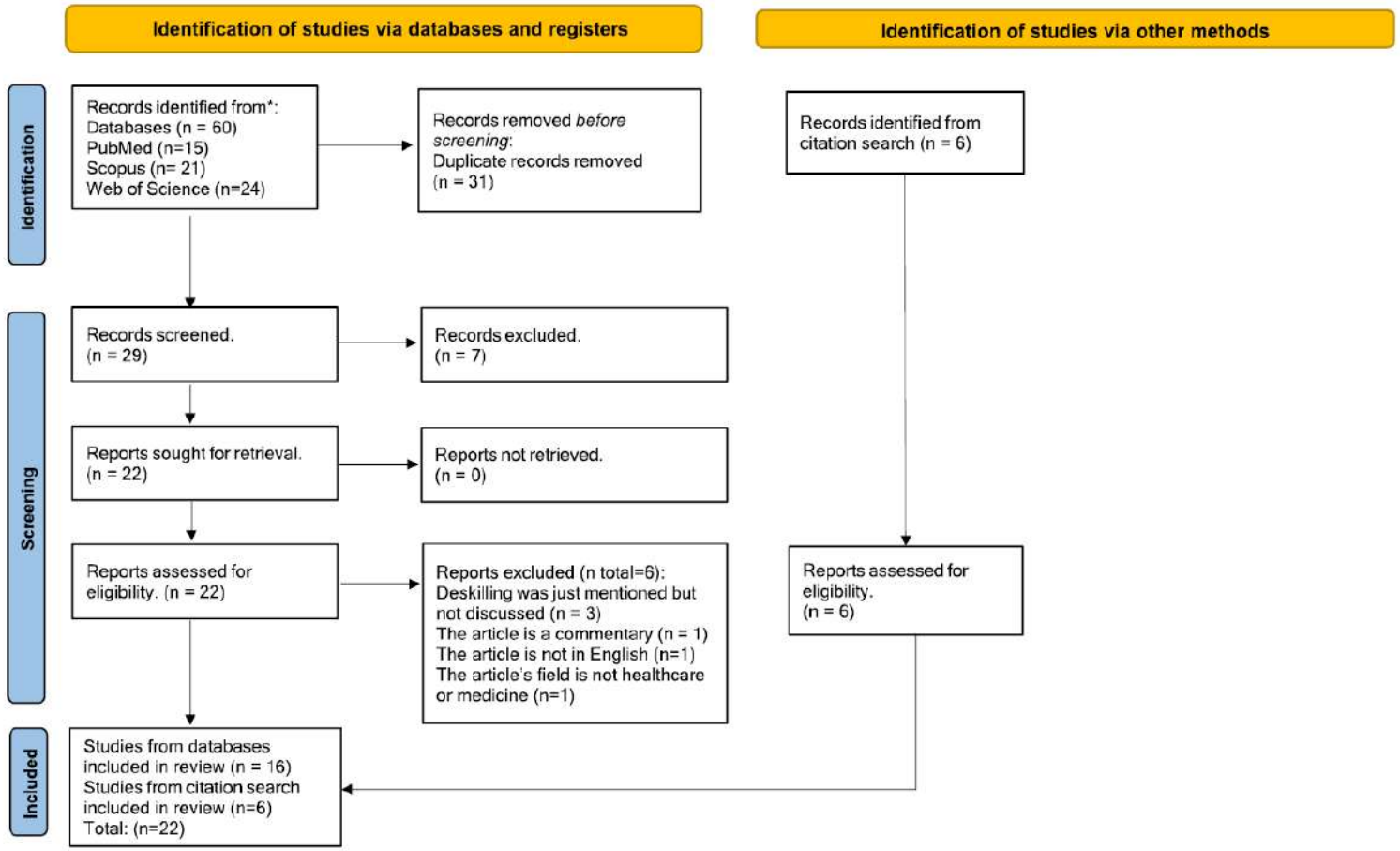
Home › Artificial Intelligence Review › Article

AI-induced Deskilling in Medicine: A Mixed-Method Review and Research Agenda for Healthcare and Beyond

4 mesi fa…

Home › Artificial Intelligence Review › Article

AI-induced Deskilling in Medicine: A Mixed-Method Review and Research Agenda for Healthcare and Beyond

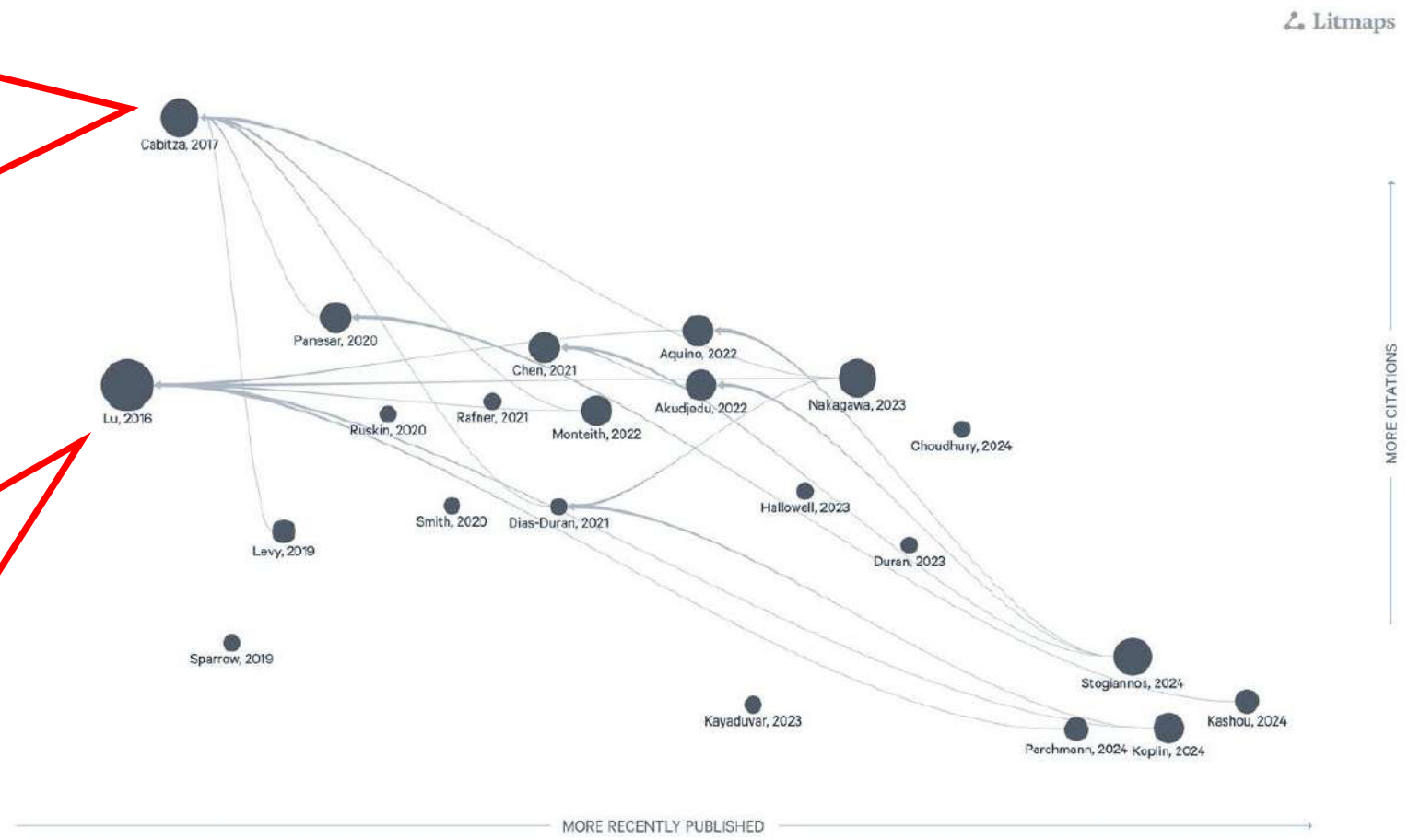


Home › Artificial Intelligence Review › Article

AI-induced Deskilling in Medicine: A Mixed-Method Review and Research Agenda for Healthcare and Beyond

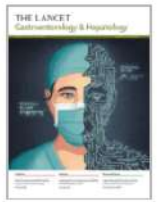
Cabitza F, Rasoini R, Gensini GF (2017) Unintended consequences of machine learning in medicine. JAMA 318(6):517–518

Lu J (2016) Will medical technology deskill doctors? Int Educ Stud 9(7):130–134



The Lancet Gastroenterology & Hepatology

Volume 10, Issue 10, October 2025, Pages 896-903

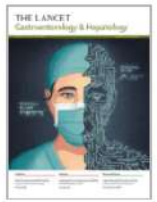


ARTICLES · Volume 10, Issue 10, P896-903, October 2025

Endoscopist deskilling risk after exposure to artificial intelligence in colonoscopy: a multicentre, observational study

Krzysztof Budzyń, MD [a,b] · Marcin Romańczyk, MD [a,b] · Diana Kitala, PhD [c] · Paweł Kołodziej, MD [d] · Marek Bugajski, MD [e] · Hans O Adami, MD [f,g] · et al. Show more

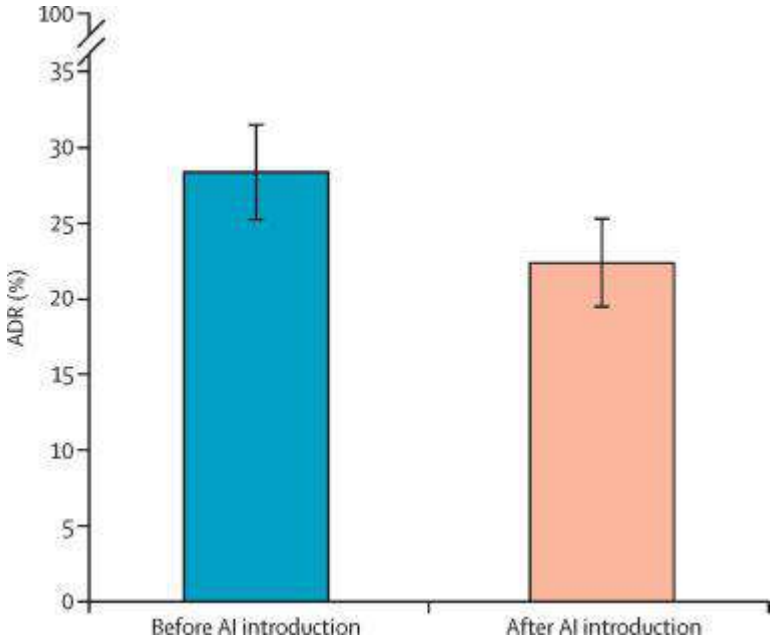
The Lancet Gastroenterology & Hepatology
Volume 10, Issue 10, October 2025, Pages 896-903



ARTICLES · Volume 10, Issue 10, P896-903, October 2025

Endoscopist deskilling risk after exposure to artificial intelligence in colonoscopy: a multicentre, observational study

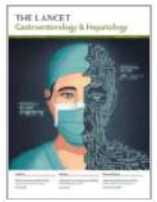
Krzysztof Budzyń, MD [a,b] · Marcin Romańczyk, MD [a,b] · Diana Kitala, PhD [c] · Paweł Kołodziej, MD [d] · Marek Bugajski, MD [e] · Hans O Adami, MD [f,g] · et al. Show more



Findings

Between Sept 8, 2021, and March 9, 2022, 1443 patients underwent non-AI assisted colonoscopy before (n=795) and after (n=648) the introduction of AI (median age 61 years [IQR 45–70], 847 [58·7%] female, 596 [41·3%] male). The ADR of standard colonoscopy decreased significantly from 28·4% (226 of 795) before to 22·4% (145 of 648) after exposure to AI, corresponding with an absolute difference of −6·0% (95% CI −10·5 to −1·6; p=0·0089). In multivariable logistic regression analysis, exposure to AI

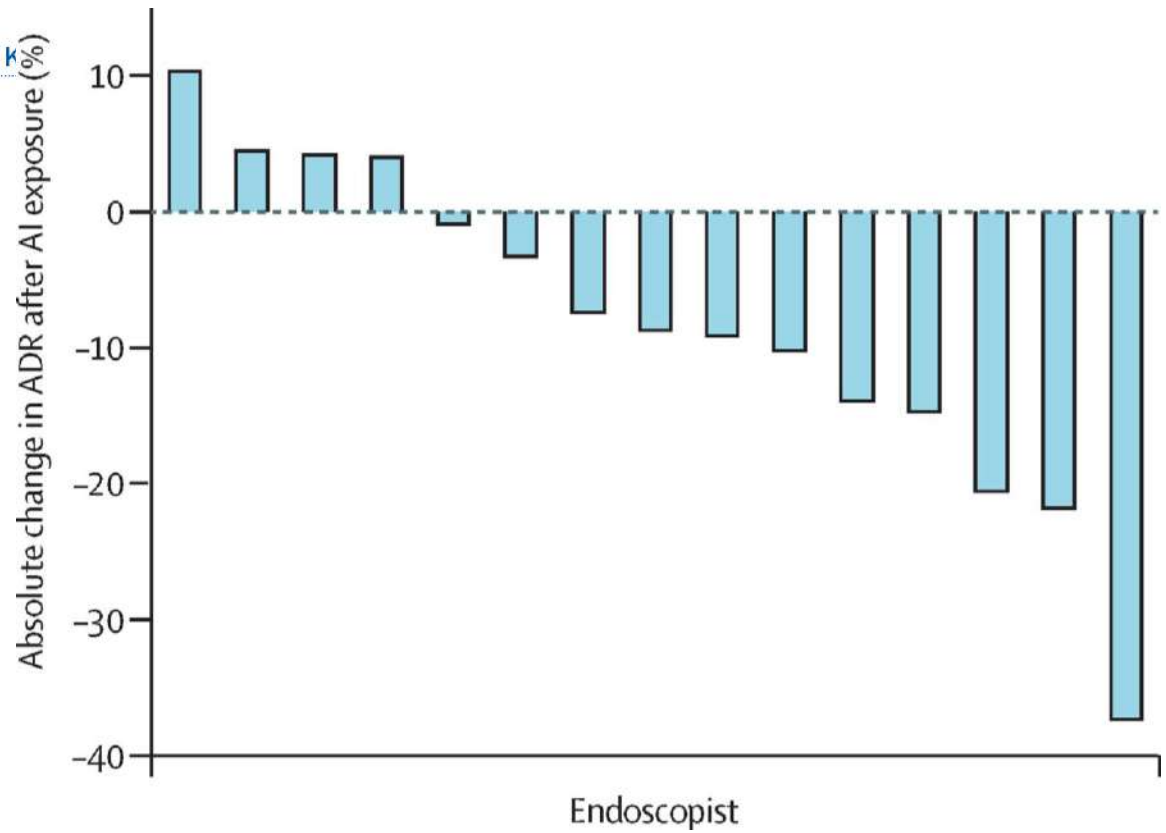
The Lancet Gastroenterology & Hepatology
Volume 10, Issue 10, October 2025, Pages 896-903



ARTICLES · Volume 10, Issue 10, P896-903, October 2025

Endoscopist deskilling risk after exposure to artificial intelligence in colonoscopy: a multicentre, observational study

Krzysztof Budzyń, MD [a,b] · Marcin Romańczyk, MD [a,b] · Diana Kitala, PhD [c] · Paweł K
· Hans O Adami, MD [f,g] · et al. Show more



Home › Artificial Intelligence Review › Article

AI-induced Deskilling in Medicine: A Mixed-Method Review and Research Agenda for Healthcare and Beyond

Deskilling (Erosione delle Competenze)

- **Definizione**: La degradazione o la perdita di competenze professionali precedentemente acquisite, tipicamente a causa di una pratica ridotta dovuta alla sostituzione tecnologica.
- **Manifestazione**: I medici passano da un ruolo di decisori attivi a supervisori di raccomandazioni generate dall'IA. Questo riduce la competenza tecnica e il giudizio clinico sfumato.
- **Impatto**: Perdita di fiducia professionale e autonomia, rafforzando la dipendenza dall'IA.

Upskilling Inhibition (Inibizione dello Sviluppo)

- **Definizione**: L'ostacolo o la soppressione delle opportunità di acquisire nuove competenze, specialmente per i professionisti in formazione.
- **Manifestazione**: I tirocinanti sono meno esposti a casi difficili e a scenari decisionali complessi, poiché l'IA fornisce costantemente le soluzioni.
- **Impatto**: Una generazione di clinici meno preparata a operare senza assistenza, aggravando la dipendenza sistemica a lungo termine.

Home › Artificial Intelligence Review › Article

AI-induced Deskilling in Medicine: A Mixed-Method Review and Research Agenda for Healthcare and Beyond

Mappatura delle Vulnerabilità: Risultati della Revisione Sistematica

L'analisi di 22 studi ha rivelato 17 preoccupazioni specifiche, classificate secondo 7 domini di competenza. Le aree più colpite sono quelle che richiedono un'interazione umana sfumata e un ragionamento complesso.

Esame Fisico e Identificazione dei Segni Fisici

Preoccupazioni: Deterioramento delle capacità di esame fisico, perdita di abilità manuali.

Giudizio Clinico

Preoccupazioni: Peggioramento del ragionamento clinico, riluttanza a fornire valutazioni definitive, processo decisionale più povero.

Diagnosi Differenziale

Preoccupazioni: Diminuzione dell'accuratezza diagnostica, riduzione delle conoscenze cliniche e dell'interpretazione.

Comunicazione Clinica e Gestione delle Preoccupazioni del Paziente

Preoccupazioni: Peggioramento della comunicazione medico-paziente, deterioramento della relazione.

Competenze Specifiche dell'IA

Preoccupazioni: Dipendenza dagli strumenti, incapacità di comprendere e contestare gli output dell'IA, errori medici da HCI.

Competenze Organizzative

Preoccupazioni: Impatto negativo sull'integrità del sistema, impatto negativo sul lavoro di squadra.

Mantenimento del Benessere del Paziente

Preoccupazioni: Deterioramento delle competenze morali, compromissione della sicurezza del paziente.

Home › Artificial Intelligence Review › Article

AI-induced Deskilling in Medicine: A Mixed-Method Review and Research Agenda for Healthcare and Beyond

Home › Artificial Intelligence Review › Article

AI-induced Deskilling in Medicine: A Mixed-Method Review and Research Agenda for Healthcare and Beyond

Distinguere i Fenomeni: Differenziare chiaramente tra *deskilling* (perdita di competenze), *upskilling inhibition* (mancata acquisizione) e *re-skilling* (trasformazione delle competenze). Non ogni cambiamento è una perdita.

Monitoraggio Longitudinale in Situ: Andare oltre gli studi una tantum. È necessario monitorare l'interazione clinico-IA nel tempo e in contesti reali (es. reparti ospedalieri) per mappare la traiettoria dell'impatto sulle competenze.

Approcci Qualitativi e Misti: Le metriche quantitative (come l'accuratezza) non bastano. Sono necessarie interviste, focus group e studi di caso per comprendere l'impatto su identità professionale, fiducia e autonomia.

Mitigare il Publication Bias: Dare priorità all'analisi dei fallimenti ("failure analyses") e alla segnalazione trasparente dei casi in cui l'IA ha contribuito a errori o a degrado delle competenze, per ottenere una visione equilibrata.

Home › Artificial Intelligence Review › Article

AI-induced Deskilling in Medicine: A Mixed-Method Review and Research Agenda for Healthcare and Beyond

Sviluppare l'Intelligenza Ibrida

Creare strumenti di IA che funzionino come partner interattivi, non come oracoli. I sistemi dovrebbero fornire ragionamenti trasparenti (Explainable AI) o persino porre domande "provocatorie" per stimolare il pensiero critico del clinico.

Introdurre "Attrito" Costruttivo

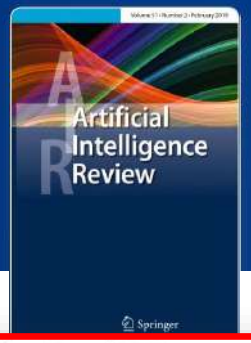
Progettare interazioni che richiedano un impegno attivo. Ad esempio, un sistema potrebbe astenersi intenzionalmente dal dare un'interpretazione definitiva in alcuni casi, spingendo il clinico a formulare la propria analisi prima di rivelare il suggerimento dell'IA.

Stabilire Protocolli di Collaborazione Uomo-IA

Invece di un uso ad-hoc, definire linee guida chiare. Un esempio è il "double-reading" in radiologia, dove sia il radiologo che l'IA interpretano un'immagine in modo indipendente prima di confrontare i risultati, mantenendo alta la vigilanza diagnostica.

Home › Artificial Intelligence Review › Article

AI-induced Deskilling in Medicine: A Mixed-Method Review and Research Agenda for Healthcare and Beyond



Hierarchy of evidence for AI and XAI empirical studies.

Level 1 (strongest)	Meta-analyses and systematic reviews of randomized controlled trials or experimental studies involving real practitioners
Level 2	Single experimental study (randomized, controlled) with prospective real-world cases considered by real practitioners in real-world settings.
Level 3	Single quasi-experimental study (e.g., nonrandomized, with concurrent or historical controls) involving prospective real-world cases considered by real practitioners in real-world settings.
Level 4	Single experimental study (randomized, controlled) with retrospective real-world cases considered by real practitioners in simulated/laboratory settings.
Level 5	Single quasi-experimental study (e.g., nonrandomized, with concurrent or historical controls) involving retrospective real-world (or simulated) cases considered by real practitioners in simulated/laboratory settings
Level 6	Single quasi-experimental study (e.g., nonrandomized, with concurrent or historical controls) involving simulated cases considered by human participants but not real practitioners in laboratory settings
Level 7	Supervised machine learning train/test studies with external validation (multiple datasets in longitudinal or cross-section/multi-site settings)
Level 8	Supervised machine learning train/test studies with internal validation
Level 9	Consensus opinions of authoritative bodies (e.g., nationally recognized guideline groups with robust peer review processes, notified bodies, standardization organizations)
Level 10 (weakest)	Opinions of recognized experts and case studies

From: Famiglini, L., Campagner, A., Barandas, M., La Maida, G. A., Gallazzi, E., & Cabitza, F. (2024). **Evidence-based XAI**: An empirical approach to design more effective and explainable decision support systems. ***Computers in Biology and Medicine***, *170*, 108042.

Home › Artificial Intelligence Review › Article

AI-induced Deskilling in Medicine: A Mixed-Method Review and Research Agenda for Healthcare and Beyond

Aggiornare i Curricula Medici: Integrare una formazione completa sull'IA, che includa i suoi limiti, le modalità di fallimento e l'uso critico. Allo stesso tempo, rafforzare la pratica delle competenze di base senza l'ausilio dell'IA.

Promuovere Competenze Umane Complementari: Porre maggiore enfasi sullo sviluppo di abilità che l'IA non può replicare: empatia, giudizio etico, gestione della complessità e dell'incertezza, e comunicazione sfumata.

Istituire Pratiche di Mantenimento delle Competenze: Ispirandosi all'aviazione (dove i piloti devono registrare ore di volo manuale), le istituzioni sanitarie potrebbero richiedere ai clinici di eseguire periodicamente esercizi di ragionamento clinico senza supporto algoritmico.

Imparare da Altri Settori ad Alto Rischio: Adattare strategie e principi di governance da campi come l'aviazione, il diritto e la finanza, che affrontano sfide simili.

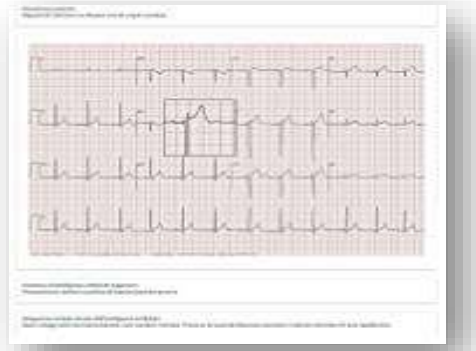
Home › Artificial Intelligence Review › Article

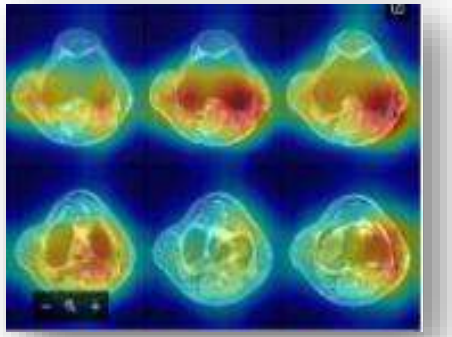
AI-induced Deskilling in Medicine: A Mixed-Method Review and Research Agenda for Healthcare and Beyond

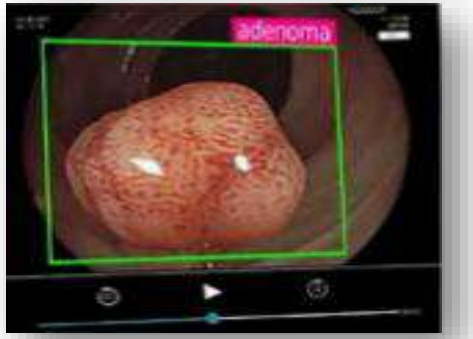
Home › Artificial Intelligence Review › Article

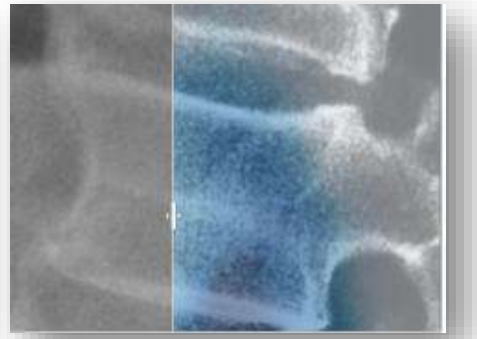
AI-induced Deskilling in Medicine: A Mixed-Method Review and Research Agenda for Healthcare and Beyond

?

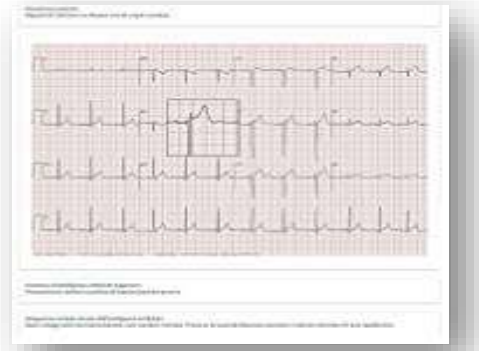


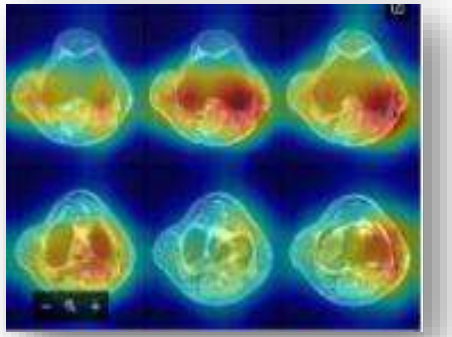


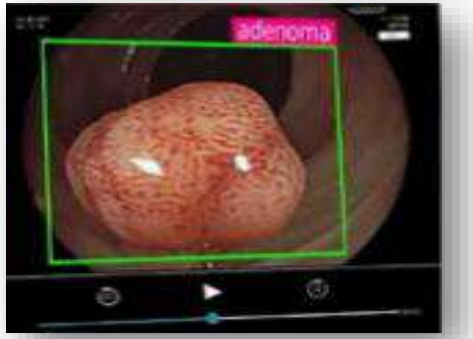


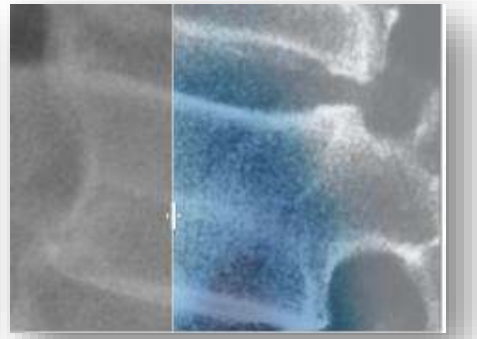


Negli ultimi 3 anni abbiamo condotto studi in 6 setting diagnostici, 5 modalities (enteroscopia con videocapsula, colonscopia, raggi X per fratture vertebrali, ECG per battiti patologici, MRI per lesioni al ginocchio).

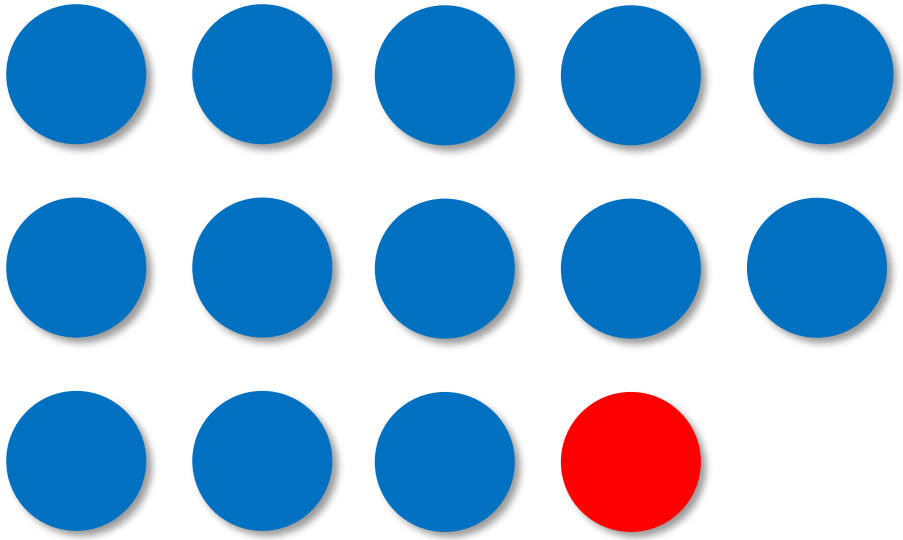


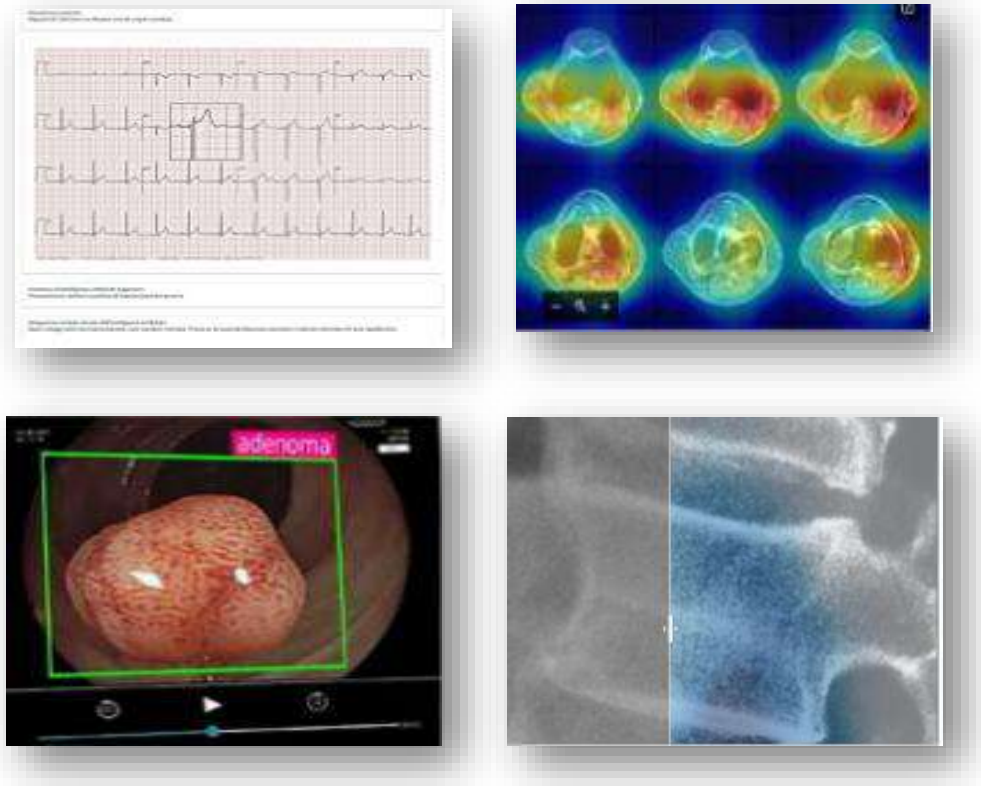




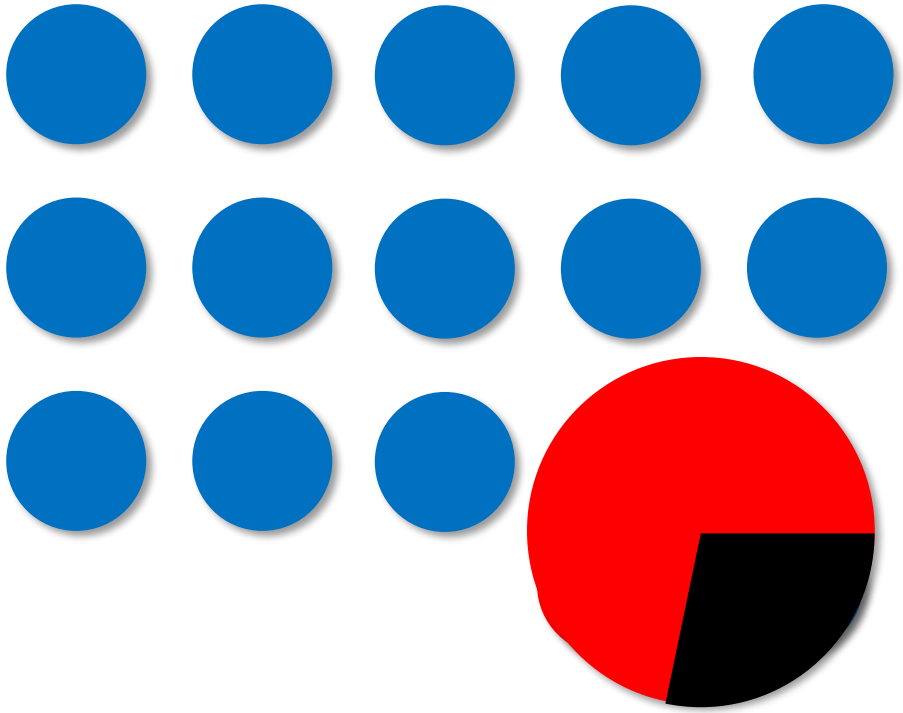


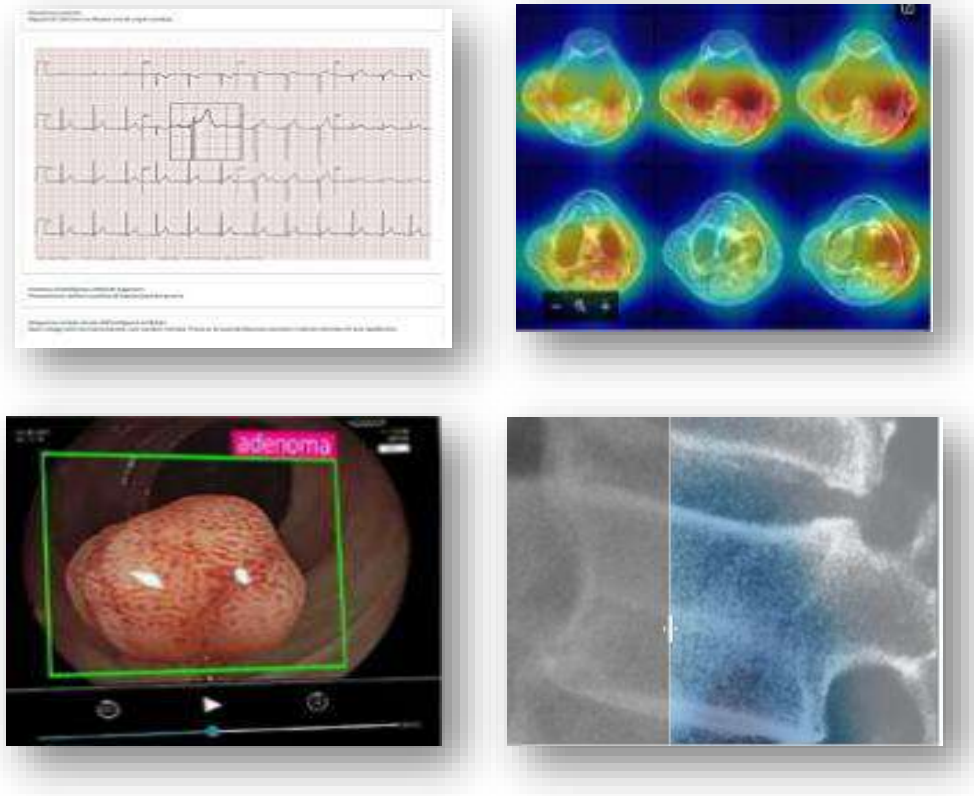
Abbiamo osservato il supporto di vari tipi di IA in 16641 casi diagnostici, coinvolgendo 330 medici.



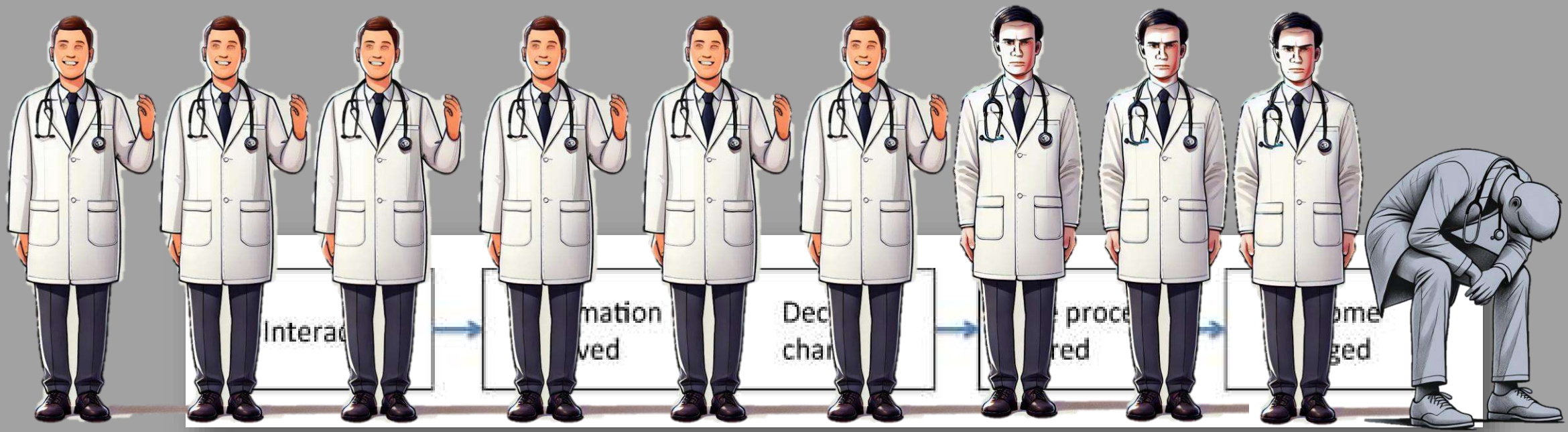


Nel 7% dei casi la IA ha fatto cambiare idea al medico.





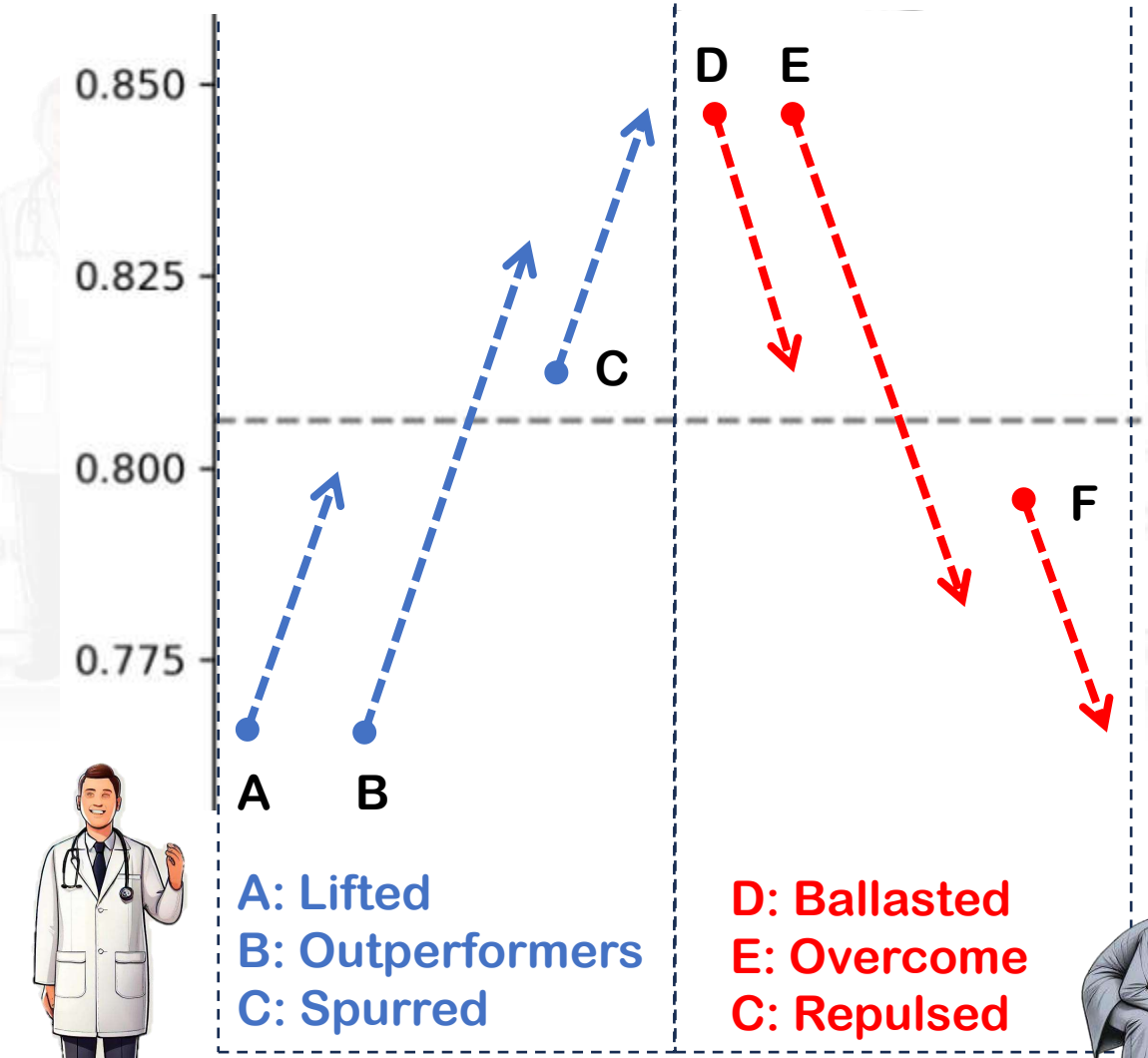
In poco più di 1/4
facendolo sbagliare!
Automation bias!



Quasi il 60% dei medici è stato "amplificato", ma più di un decimo danneggiato!

57% 11%

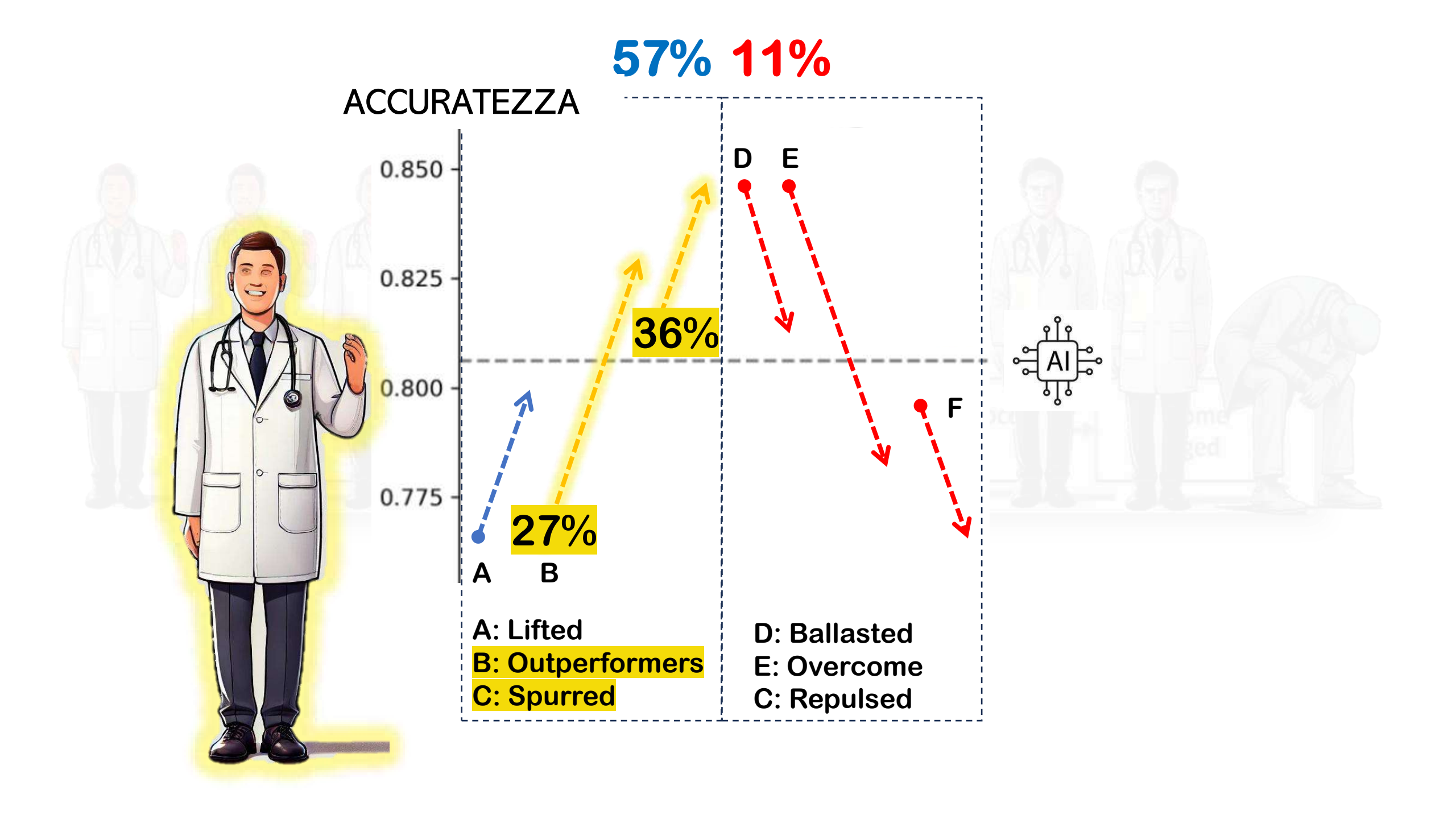
ACCURATEZZA

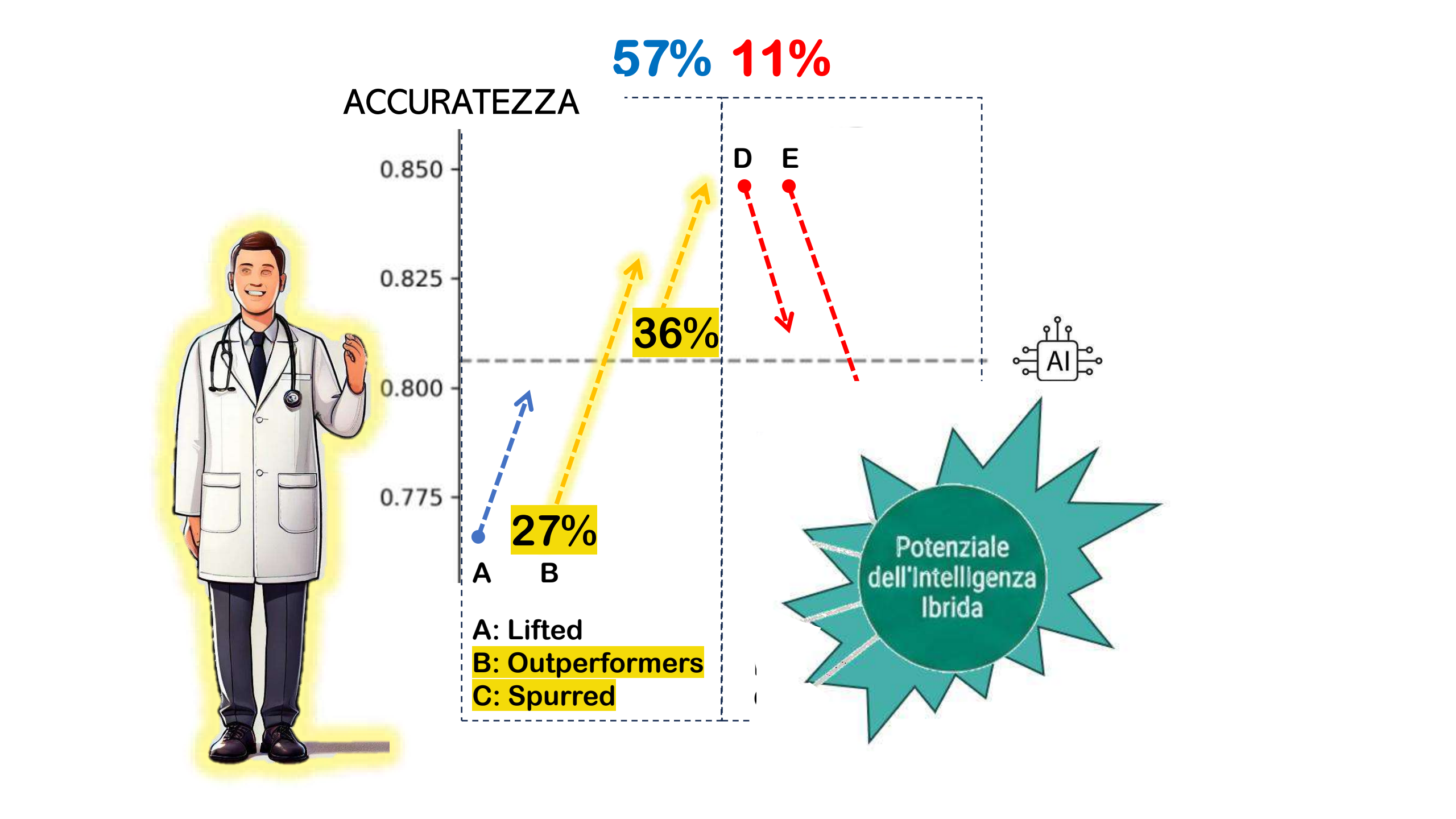


A: Lifted
B: Outperformers
C: Spurred

D: Ballasted
E: Overcome
C: Repulsed

Unaffected: 33%
un terzo non cambia


57% 11%
ACCURATEZZA
0.850
0.825
0.800
0.775
D
E
F
36%
27%
A
B
AI
A: Lifted
B: Outperformers
C: Spurred
D: Ballasted
E: Overcome
C: Repulsed


57% 11%
ACCURATEZZA
0.850
0.825
0.800
0.775
D
E
36%
27%
A
B
A: Lifted
B: Outperformers
C: Spurred
AI
Potenziale dell'Intelligenza Ibrida

TECHNO-STRESS

Techno-Stress

Stress AI-correlato

TECHNO-STRESS

STRESS AI-CORRELATO

TECHNO-STRESS

STRESS AI-CORRELATO

SKILLS

CONTROL

AGENCY

COOPERATION

ISO 45003 — Occupational health and safety management — Psychological health and safety at work — Guidelines for managing psychosocial risks

World Health Organization

International Labour Organization

European Agency for Safety and Health at Work

TECHNO-STRESS

STRESS AI-CORRELATO

SKILLS	SAPERE
CONTROL	COMPRENDERE
AGENCY	AGIRE
COOPERATION	INTERAGIRE COLLABORARE

ISO 45003 Occupational health and safety management — Psychological health and safety at work — Guidelines for managing psychosocial risks

World Health Organization

International Labour Organization

European Agency for Safety and Health at Work

TECHNO-STRESS

STRESS AI-CORRELATO

SKILLS

CONTROL

AGENCY

COOPERATION

1.1	**Complessità e competenza percepita** (*techno-complexity*)	Difficoltà nel comprendere o utilizzare sistemi complessi e in continua evoluzione; genera senso di inadeguatezza e fatica cognitiva.
1.2	**Deskilling creativo e dipendenza cognitiva**	Progressiva perdita di abilità critiche e di padronanza professionale per eccessiva delega all'IA; rischio di impoverimento del sapere esperto individuale e collettivo.
2.1	**Opacità e attribuzione di responsabilità** (*black-box stress*)	Impossibilità di comprendere le logiche decisionali dell'IA e difficoltà nell'attribuzione delle responsabilità; genera sfiducia e senso di impotenza.
2.2	**Affidabilità degli strumenti** (*tool reliability stress*)	Guasti, latenza o incoerenza degli output compromettono la fiducia e la capacità di supervisione, riducendo il senso di controllo operativo.
3.1	**Sorveglianza algoritmica e perdita di autonomia**	Controllo granulare delle prestazioni e prescrittività delle metriche che limitano la libertà d'azione e la responsabilità personale.
3.2	**Sovraccarico digitale** (*techno-overload / workslop*)	Eccesso di input, notifiche e richieste che impongono ritmi insostenibili, riducendo la capacità di giudizio e di decisione autonoma.
3.3	**Insicurezza occupazionale** (*techno-insecurity*)	Timore di sostituzione o dequalificazione del ruolo dovuta all'automazione e all'introduzione dell'IA; induce passività e conformismo.
4.1	**Profilazione predittiva e stigma**	Assegnazione di punteggi o etichette probabilistiche che influenzano reputazione e opportunità, generando ansia, percezione d'ingiustizia e sfiducia nei sistemi.
4.2	**Conflitto etico/morale** (*moral distress*)	Disagio psicologico dovuto a decisioni o raccomandazioni dell'IA percepite come contrarie ai valori etici o deontologici della professione.
4.3	**Invasione/confusione dei confini** (*telepressure / work–home conflict*)	Erosione del confine tra lavoro e vita privata a causa dell'iperconnessione e della reperibilità costante; riduce il recupero e la qualità relazionale.
4.4	**Sicurezza informatica e privacy** (*security/privacy anxiety*)	Paura di errori, leak o monitoraggio costante; compromette la fiducia nei sistemi e nei processi collaborativi.
4.5	**Incertezza tecnologica e adattamento continuo** (*techno-uncertainty*)	Evoluzione imprevedibile di modelli, policy e procedure che ostacola la stabilità operativa e la costruzione di fiducia reciproca.
4.6	**Isolamento e riduzione del supporto sociale**	Diminuzione delle interazioni umane significative, feedback automatizzati e perdita di senso di appartenenza nei contesti mediati dall'IA.

ISO 45003 — Occupational health and safety management — Psychological health and safety at work — Guidelines for managing psychosocial risks

World Health Organization

International Labour Organization

European Agency for Safety and Health at Work

TECHNO-STRESS

STRESS AI-CORRELATO

SKILLS

CONTROL

AGENCY

COOPERATION

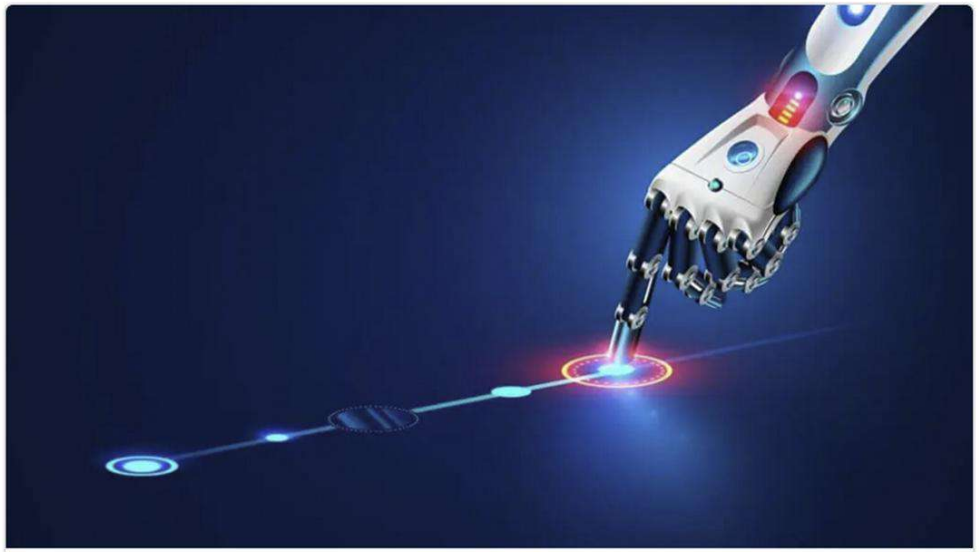
1.1	**Complessità e competenza percepita** (*techno-complexity*)	Difficoltà nel comprendere o utilizzare sistemi complessi e in continua evoluzione; genera senso di inadeguatezza e fatica cognitiva.
1.2	**Deskilling creativo e dipendenza cognitiva**	Progressiva perdita di abilità critiche e di padronanza professionale per eccessiva delega all'IA; rischio di impoverimento del sapere esperto individuale e collettivo.
2.1	**Opacità e attribuzione di responsabilità** (*black-box stress*)	Impossibilità di comprendere le logiche decisionali dell'IA e difficoltà nell'attribuzione delle responsabilità; genera sfiducia e senso di impotenza.
2.2	**Affidabilità degli strumenti** (*tool reliability stress*)	Guasti, latenza o incoerenza degli output compromettono la fiducia e la capacità di supervisione, riducendo il senso di controllo operativo.
3.1	**Sorveglianza algoritmica e perdita di autonomia**	Controllo granulare delle prestazioni e prescrittività delle metriche che limitano la libertà d'azione e la responsabilità personale.
3.2	**Sovraccarico digitale** (*techno-overload / workslop*)	Eccesso di input, notifiche e richieste che impongono ritmi insostenibili, riducendo la capacità di giudizio e di decisione autonoma.
3.3	**Insicurezza occupazionale** (*techno-insecurity*)	Timore di sostituzione o dequalificazione del ruolo dovuta all'automazione e all'introduzione dell'IA; induce passività e conformismo.
4.1	**Profilazione predittiva e stigma**	Assegnazione di punteggi o etichette probabilistiche che influenzano reputazione e opportunità, generando ansia, percezione d'ingiustizia e sfiducia nei sistemi.
4.2	**Conflitto etico/morale** (*moral distress*)	Disagio psicologico dovuto a decisioni o raccomandazioni dell'IA percepite come contrarie ai valori etici o deontologici della professione.
4.3	**Invasione/confusione dei confini** (*telepressure / work–home conflict*)	Erosione del confine tra lavoro e vita privata a causa dell'iperconnessione e della reperibilità costante; riduce il recupero e la qualità relazionale.
4.4	**Sicurezza informatica e privacy** (*security/privacy anxiety*)	Paura di errori, leak o monitoraggio costante; compromette la fiducia nei sistemi e nei processi collaborativi.
4.5	**Incertezza tecnologica e adattamento continuo** (*techno-uncertainty*)	Evoluzione imprevedibile di modelli, policy e procedure che ostacola la stabilità operativa e la costruzione di fiducia reciproca.
4.6	**Isolamento e riduzione del supporto sociale**	Diminuzione delle interazioni umane significative, feedback automatizzati e perdita di senso di appartenenza nei contesti mediati dall'IA.

ISO 45003 Occupational health and safety management — Psychological health and safety at work — Guidelines for managing psychosocial risks

World Health Organization

International Labour Organization

European Agency for Safety and Health at Work



Agenda Digitale.eu

AI e sicurezza sul lavoro: il nuovo mandato del Servizio di Prevenzione e Protezione

05 Nov 2025

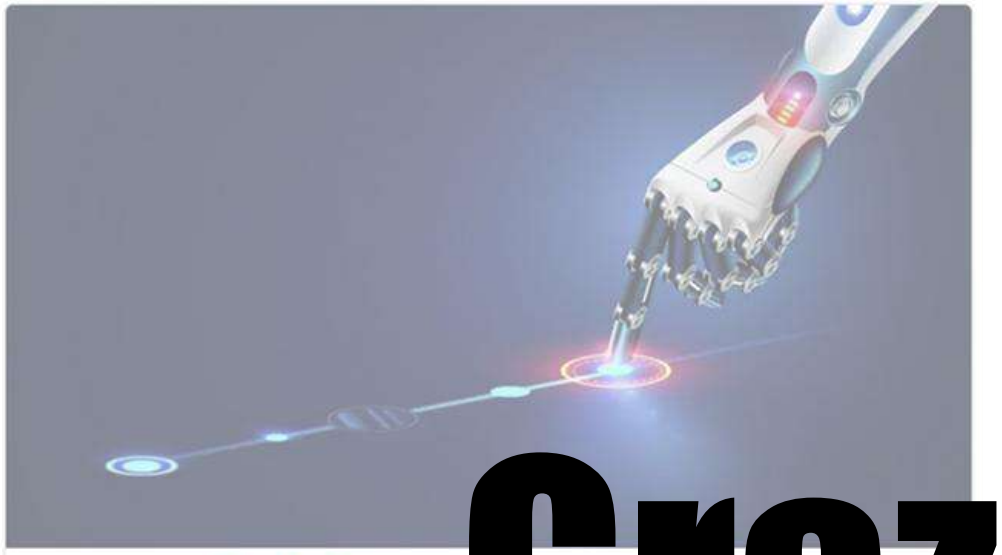
di **Federico Cabitza**

Agenda Digitale.eu

L'Algoritmo in Azienda: perché la governance dell'IA è (soprattutto) una questione di salute e sicurezza

Dicembre 2025

di **Federico Cabitza**



Agenda Digitale.eu

AI e sicurezza sul la[illegible] mandato del Servizio di Prevenzione e Protezione

05 Nov 2025

di **Federico Cabitza**

Agenda Digitale.eu

[illegible]goritmo in Azienda: perché la governan[illegible] questio[illegible]

Dicembre 2[illegible]

di **Federic[illegible]**

Grazie!

federico.cabitza@unimib.it

@cabitzaf

federicocabitza

federicocabitza.net