



UNIMORE

UNIVERSITÀ DEGLI STUDI DI
MODENA E REGGIO EMILIA

Dipartimento di Scienze Biomediche,
Metaboliche e Neuroscienze

**Corso di Laurea Magistrale in
Scienze Infermieristiche ed Ostetriche**

Presidente: Prof.ssa Alessia Cadamuro

**“Efficacia delle soluzioni di intelligenza artificiale
per l’ottimizzazione delle risorse e la valutazione
degli esiti in ambito ambulatoriale e chirurgico:
Revisione narrativa della letteratura con
stratificazione per tecnologia, compito e outcome”**

Relatore:

Prof. Marco Foracchia

Studente:

Giangiuseppe Scopelliti

Anno Accademico 2024-2025

Indice

CAPITOLO 1 - Introduzione	3
1.1 Pressione sulle risorse ambulatoriali e chirurgiche e liste d’attesa: il contesto attuale . 3	
1.2 Opportunità dell’intelligenza artificiale per l’ottimizzazione delle risorse sanitarie 5	
1.3 Obiettivo della tesi e domande di ricerca	7
CAPITOLO 2 - Contesto e quadro regolatorio	10
2.1 Definizioni e concetti fondamentali	11
2.2 Norme e guidance su IA in ambito sanitario	23
2.2.1 Guidance dell’Organizzazione Mondiale della Sanità: etica e regolamentazione . 23	
2.2.2 Il quadro normativo europeo: l’AI Act (Regolamento UE 2024/1689) e il settore sanitario.....	32
2.2.3 La legge italiana 23 settembre 2025 n. 132: principi e implicazioni per l’uso dell’IA in sanità	40
2.3 Implicazioni operative per i professionisti sanitari.....	46
CAPITOLO 3 - Metodi della Revisione	50
3.1 Disegno dello studio e formulazione del PICO	50
3.2 Ricerca su basi di dati tradizionali (PubMed)	50
3.3 Piattaforme di Intelligenza Artificiale per la ricerca (Consensus, Elicit, ChatGPT)	53
3.4 Gestione delle fonti e screening.....	54
CAPITOLO 4 - Risultati e Stratificazioni principali	60
4.1 Per ambiente di cura	60
4.1.1 Ambulatoriale/Outpatient (no show, triage/overbooking, capacity planning).....	60
4.1.2 Chirurgico/OR (stima durate, block scheduling, turnover, cancellazioni)	62
4.2 Per tecnologia IA	67
4.2.1 ML supervisionato (gradient boosting, random forest, prediction di no show e durate).....	67
4.2.2 Deep Learning (RNN/LSTM, CNN – serie temporali di sala/attese)	69
4.2.3 Reinforcement Learning - politiche di scheduling adattivo.....	71
4.2.4 Ibridi ML + Ricerca Operativa/Programmazione matematica (cluster predict & schedule, ottimizzazione vincolata)	73
4.3 Per compito/uso.....	76
4.3.1 Previsione (durata interventi, no show, lunghezza lista)	76
4.3.2 Ottimizzazione (allocazione slot/risorse, block management, riorganizzazione agende)	79
4.3.3 Supporto decisionale operativo (Operational Decision Support Systems – ODSS): simulazioni “what-if” e dashboard.....	82
4.4 Per outcome.....	84

4.4.1 Efficienza (tempi di attesa, utilizzo sale/ambulatori, throughput, turnover, cancellazioni).....	84
4.4.2 Qualità/continuità (aderenza agli appuntamenti, riprogrammazioni, no show) ..	87
4.4.3 Costi/risorse (risparmio ore-persona, blocchi inutilizzati)	90
4.4.4 Sicurezza/equità (bias, trasparenza, accettazione staff/pazienti)	92
CAPITOLO 5 - Discussione	96
5.1 Sintesi delle evidenze per ciascuna stratificazione	96
5.1.1 Per ambiente di cura (ambulatoriale vs chirurgico)	96
5.1.2 Per tecnologia IA (ML, DL, RL, ibridi).....	98
5.1.3 Per compito/uso (previsione, ottimizzazione, supporto decisionale).....	100
5.1.4 Per outcome (efficienza, qualità, costi, equità/sicurezza)	103
5.2 Fattori critici per l'implementazione	108
5.2.1 Qualità e disponibilità dei dati	108
5.2.2 Integrazione nei sistemi informativi	110
5.2.3 Governance e accettabilità.....	112
5.2.4 Change management e formazione del personale	114
5.3 Stato di sviluppo delle soluzioni identificate	117
5.3.1 Studi di simulazione o proof-of-concept.....	118
5.3.2 Applicazioni sperimentate in contesti reali ma non ancora routinari	119
5.3.3 Soluzioni già integrate nella pratica clinico-organizzativa quotidiana.....	122
5.4 Principali gap di ricerca	124
CAPITOLO 6 - Limiti della revisione	127
6.1 Limiti legati al disegno della revisione	127
6.2 Eterogeneità degli studi inclusi e degli outcome analizzati.....	128
6.3 Limiti legati alla qualità dei dati e alle evidenze disponibili	129
6.4 Limitata evidenza sugli esiti clinici e sulla sostenibilità nel tempo.....	130
6.5 Rapidità di evoluzione tecnologica e rischio di obsolescenza delle evidenze.....	132
CAPITOLO 7 - Conclusioni	134
Bibliografia	137
Appendice A - Dichiarazione di utilizzo di strumenti basati su Intelligenza Artificiale	143
Allegato A - Griglia estrazione dati degli studi	

Keywords (Mesh): Artificial Intelligence; Waiting Lists; Ambulatory Care; Operating Rooms; Resource Allocation.

CAPITOLO 1 - Introduzione

1.1 Pressione sulle risorse ambulatoriali e chirurgiche e liste d'attesa: il contesto attuale

Le lunghe liste d'attesa per visite specialistiche, esami diagnostici e interventi chirurgici rappresentano da anni una delle principali criticità dei sistemi sanitari pubblici a livello internazionale (Fatebenefratelli, 2023). In Italia, in particolare, il problema delle attese ha assunto una rilevanza allarmante: quasi la metà dei cittadini (49,5%) individua nelle liste d'attesa la principale barriera di accesso alle prestazioni del Servizio Sanitario Nazionale (Fatebenefratelli, 2023). Le segnalazioni raccolte indicano che le attese più critiche riguardano soprattutto le prime visite ambulatoriali (25,3% delle segnalazioni), gli esami diagnostici (19,2%) e i ricoveri chirurgici programmati (13,3%) (Fatebenefratelli, 2023).

Distribuzione delle principali criticità segnalate dai cittadini in relazione alle liste d'attesa
(Fatebenefratelli, 2023)

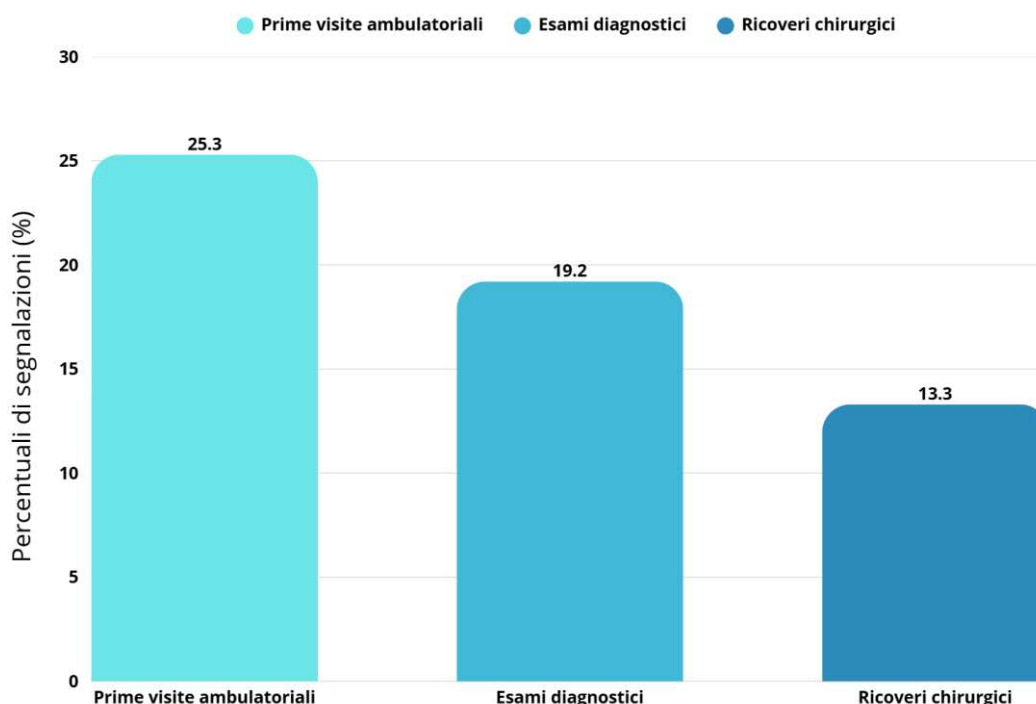


Figura 1 - I dati mostrano la distribuzione percentuale delle segnalazioni dei cittadini relative ai tempi di attesa, con particolare incidenza sulle prime visite ambulatoriali, sugli esami diagnostici e sui ricoveri chirurgici programmati. Fonte: Fatebenefratelli (2023).

La situazione ha spinto un numero crescente di persone a rivolgersi a strutture private a pagamento o addirittura a rinunciare alle cure: secondo i dati ISTAT, nel

2022 circa il 7% della popolazione (*oltre 4 milioni di persone*) ha dovuto rinunciare a visite specialistiche necessarie, nella maggior parte dei casi proprio a causa di liste d'attesa eccessivamente lunghe (*Pagella Politica, 2024*). Tali percentuali, sebbene leggermente migliorate rispetto al periodo dell'emergenza COVID-19, restano superiori ai livelli pre-pandemici (era il 6,3% nel 2019) e testimoniano una perdurante difficoltà di accesso tempestivo alle cure (*Pagella Politica, 2024*). La pandemia ha infatti aggravato gli arretrati: ancora nel primo semestre 2023, molte Regioni non erano riuscite a ripristinare i volumi di prestazioni erogate prima del 2020, evidenziando criticità nel recupero dell'offerta assistenziale pre-Covid (*AGENAS, 2023*).

Le cause delle lunghe attese sono molteplici e interconnesse: l'invecchiamento demografico e l'aumento della domanda di prestazioni sanitarie, una programmazione e allocazione delle risorse non sempre ottimale, carenze di personale e vincoli di bilancio, oltre agli effetti straordinari della crisi pandemica. Negli ultimi anni istituzioni nazionali e regionali hanno adottato misure per fronteggiare quella che è ormai percepita come un'"*emergenza*" del sistema sanitario (*Cittadinanzattiva, 2023*). Ad esempio, la Regione Emilia-Romagna già nel 2017 ha varato un programma straordinario per monitorare e ridurre i tempi di attesa dei ricoveri chirurgici elettivi (*Regione Emilia-Romagna, 2017*). A livello nazionale, è stato introdotto il **Piano Nazionale di Governo delle Liste di Attesa (PNGLA)**, affiancato nel 2024 da un decreto-legge specifico che ha istituito una *Piattaforma Nazionale* per la trasparenza dei tempi di attesa (*affidata ad AGENAS*) e altre misure di potenziamento dell'offerta (*Ministero della Salute, 2024*). Il Ministro della Salute ha definito "*intollerabile*" l'aumento delle liste d'attesa registrato negli ultimi anni e ha sottolineato la necessità di interventi strutturali per tutelare il diritto alla salute dei cittadini (*Ministero della Salute, 2024*). Tra le azioni intraprese vi sono il rafforzamento dei sistemi di prenotazione (CUP), l'eliminazione dal 2025 dei tetti di spesa per nuove assunzioni di personale sanitario e il ricorso a canali alternativi (*attività intra-moenia o strutture private accreditate*) qualora il servizio pubblico non riesca a garantire la prestazione entro i tempi massimi previsti. Questi provvedimenti mirano sia ad aumentare la capacità assistenziale, sia a rendere più efficiente l'utilizzo delle risorse esistenti, riducendo sprechi e inefficienze

organizzative che contribuiscono alle attese. In sintesi, la pressione sulle risorse ambulatoriali e chirurgiche e il conseguente allungamento delle liste d'attesa costituiscono oggi uno dei principali problemi per la sostenibilità e l'equità del nostro Servizio Sanitario. Ottimizzare l'uso delle risorse disponibili e trovare nuove soluzioni per migliorare l'accesso alle cure è diventata una priorità non solo per i decisori politici, ma anche per manager sanitari e professionisti sul campo.

1.2 Opportunità dell'intelligenza artificiale per l'ottimizzazione delle risorse sanitarie

In questo scenario critico, le innovazioni tecnologiche – *e in particolare l'Intelligenza Artificiale (IA)* – offrono opportunità inedite per contribuire a risolvere il problema delle liste d'attesa e migliorare l'efficienza dei servizi sanitari. L'IA è oggi riconosciuta come uno dei principali motori dell'innovazione in sanità, capace di **ridefinire i modelli di cura e assistenza** senza sostituire i professionisti ma affiancandoli quale strumento di supporto: può *amplificare le competenze cliniche, velocizzare i processi decisionali e migliorare la qualità degli interventi (AGENAS, 2025)*. Se integrata con visione strategica e responsabilità, l'IA può dare un contributo concreto nel potenziare la capacità del sistema sanitario, ad esempio riducendo le disuguaglianze territoriali nell'accesso alle cure e portando servizi avanzati anche in contesti finora svantaggiati (AGENAS, 2025). Proprio per questo, nell'ambito del Piano Nazionale di Ripresa e Resilienza (PNRR) sono previsti investimenti dedicati allo sviluppo di soluzioni di intelligenza artificiale in sanità: AGENAS è stata incaricata di realizzare una piattaforma nazionale di **IA a supporto dell'assistenza primaria e territoriale**, con l'obiettivo di migliorare l'efficacia e l'efficienza dei percorsi assistenziali (*in particolare per i pazienti cronici*) tramite strumenti di decision-support e prevenzione. Le istituzioni regionali sottolineano a loro volta come l'IA rappresenti una **"straordinaria opportunità"** per il futuro della sanità, da cogliere però con un'analisi critica delle sfide in gioco e con un'integrazione sostenibile nei processi assistenziali esistenti (*Regione Emilia-Romagna, 2024*).

Tra le possibili applicazioni dell'intelligenza artificiale in ambito organizzativo e gestionale vi sono soluzioni mirate all'ottimizzazione delle risorse ambulatoriali e delle sale operatorie. In questo contesto trovano impiego strumenti avanzati di **data**

analytics e **business intelligence**, all'interno dei quali rientrano anche modelli di **machine learning**, che consentono di analizzare grandi quantità di dati storici e identificare pattern utili alla previsione della domanda di prestazioni e delle possibili criticità organizzative.

Inoltre, algoritmi di *scheduling* basati su algoritmi di ottimizzazione tradizionale o di Intelligenza Artificiale possono aiutare a distribuire in modo ottimale i posti disponibili (*ad esempio proponendo l'overbooking controllato delle visite, oppure riorganizzando le sedute operatorie in base alla probabilità di cancellazioni e ai tempi operatori previsti*), così da **ridurre i tempi morti e aumentare il numero di pazienti trattati senza incrementare le risorse**. Le prime esperienze in questo campo indicano che l'adozione di strumenti di IA nei processi amministrativi e di allocazione delle risorse può effettivamente contribuire a *ridurre i tempi di attesa e migliorare la produttività* delle strutture sanitarie (*The Bridge & AGENAS, 2023*). Come affermato da un osservatorio promosso da AGENAS, è giunto il momento di "distillare informazioni e conoscenza dai dati sanitari, **coniugando il sapere umano con le opportunità che le nuove tecnologie di machine learning oggi permettono**". Questa integrazione tra competenze degli operatori e capacità analitiche delle macchine può abilitare un salto di qualità nella gestione delle liste d'attesa: ad esempio, consentendo di identificare in tempo reale i colli di bottiglia nel percorso del paziente, di simulare scenari alternativi ("*what-if*") per l'organizzazione di ambulatori e sale operatorie, o di supportare i decisori nel *prioritizzare* interventi e allocare personale in base ai bisogni previsti. In altri termini, l'IA si propone come strumento complementare che, se applicato correttamente, **ottimizza l'uso delle risorse sanitarie esistenti** – aumentando l'efficienza e la capacità di presa in carico – senza compromettere la qualità e la sicurezza delle cure erogate (*AGENAS, 2025*).

Le potenzialità di questa "rivoluzione" tecnologica sono già evidenti in altri ambiti medici (*basti pensare all'impiego dell'IA nella diagnostica per immagini, nella medicina di precisione o nei sistemi di supporto decisionale clinico*) e iniziano a manifestarsi anche nell'organizzazione dei servizi: numerose realtà, in Italia e all'estero, stanno sperimentando l'uso di algoritmi per rendere i flussi di lavoro più **efficienti**, riducendo il carico amministrativo e liberando tempo degli operatori a vantaggio dell'assistenza diretta ai pazienti (*4ward, 2025*). Non mancano certamente

le sfide da affrontare – dalla necessità di garantire la qualità dei dati e l'integrazione di queste soluzioni nei sistemi informativi ospedalieri, fino ai temi etici, regolatori e di accettazione da parte del personale sanitario – ma le basi per una trasformazione digitale orientata anche alla risoluzione del problema delle attese sono state gettate. L'Intelligenza Artificiale, opportunamente governata, si configura dunque come una leva strategica per aiutare il Servizio Sanitario a **ottimizzare le risorse ambulatoriali e chirurgiche**, migliorando l'accesso e la continuità di cura per i cittadini (AGENAS & Regione Emilia-Romagna, 2025).

1.3 Obiettivo della tesi e domande di ricerca

Alla luce di quanto esposto, l'obiettivo di questa tesi è valutare in che modo le soluzioni basate sull'intelligenza artificiale possano contribuire a ottimizzare le risorse in ambito ambulatoriale e chirurgico e con quali effetti sugli esiti organizzativi e assistenziali. In particolare, attraverso una **revisione narrativa della letteratura**, si intende esplorare e sintetizzare le evidenze disponibili riguardo all'applicazione di tecniche di IA nella gestione operativa di ambulatori e sale operatorie, classificando tali applicazioni in base alla tecnologia utilizzata, al tipo di compito o problema affrontato e agli outcome considerati. Le principali **domande di ricerca** a cui questo lavoro intende rispondere sono:

1. **Quali tipologie di soluzioni di intelligenza artificiale sono state sviluppate e sperimentate in ambito ambulatoriale e chirurgico per il miglioramento della pianificazione e dell'allocazione delle risorse?** In questa domanda rientrano sia le diverse **tecnologie IA** impiegate (ad es. algoritmi di *machine learning* supervisionato, modelli di *deep learning*, approcci di *reinforcement learning*, sistemi ibridi IA/ricerca operativa, ecc.), sia i **contesti specifici** di applicazione (*ambulatori specialistici, servizi diagnostici, sale operatorie, percorsi pre- e post-operatori, ecc.*). L'analisi mira a identificare le classi principali di strumenti IA adottati e il loro grado di maturità (*prototipi sperimentali vs. implementazioni reali*).
2. **Per quali finalità operative e compiti gestionali sono state utilizzate queste soluzioni di IA e come si inseriscono nei processi sanitari esistenti?** Questa

domanda approfondisce i **problemi organizzativi** affrontati dall'IA, ad esempio la previsione di eventi critici (*no-show alle visite, cancellazioni di interventi, durate degli interventi chirurgici, afflusso di pazienti*), l'**ottimizzazione delle agende** e del calendario operatorio (*scheduling di appuntamenti, assegnazione di slot o blocchi operatori, gestione delle liste di attesa e delle priorità*), nonché il **supporto decisionale** ai responsabili (*simulazione di scenari "what-if", suggerimenti per il miglioramento dei flussi, sistemi di dashboard per il monitoraggio in tempo reale delle performance*). Si indaga anche come tali soluzioni vengano integrate nei workflow sanitari e quali cambiamenti organizzativi richiedano.

3. Quali risultati ed esiti sono riportati in letteratura a seguito dell'implementazione di soluzioni di IA per l'ottimizzazione di risorse ambulatoriali e chirurgiche? In particolare, si vogliono raccogliere evidenze sull'**efficacia** di queste soluzioni in termini di indicatori quantificabili: ad esempio **riduzione dei tempi di attesa** per visite o interventi, **miglioramento dei tassi di utilizzo** di sale operatorie o ambulatori (*riduzione dei tempi inattivi, aumento del numero di pazienti trattati a parità di risorse*), diminuzione del tasso di mancata presentazione alle visite (*no-show*) o delle cancellazioni last-minute degli interventi, **ottimizzazione dei costi** o delle ore-lavoro del personale, e così via. Accanto agli esiti di efficienza, ci si propone di esaminare anche eventuali impatti su indicatori di **qualità e continuità assistenziale** (*come la soddisfazione di pazienti e operatori, la sicurezza delle cure, l'equità di accesso e la gestione di eventuali bias algoritmici*).

Attraverso la trattazione di queste domande, la tesi si propone di fornire un quadro completo e aggiornato sullo **stato dell'arte** delle applicazioni di IA volte al miglioramento dell'organizzazione sanitaria nei contesti considerati, mettendo in luce le soluzioni più promettenti, i benefici già documentati e gli eventuali limiti o gap di conoscenza. Si intende così contribuire a capire se e in che misura l'intelligenza artificiale possa rappresentare una risposta efficace alla pressione sulle risorse ambulatoriali e chirurgiche - *in termini di riduzione delle liste d'attesa e ottimizzazione dei processi* - e quali siano le condizioni chiave per

un'implementazione di successo nella pratica clinico-assistenziale. Nei capitoli successivi viene delineato il quadro teorico e normativo di riferimento relativo all'impiego dell'Intelligenza Artificiale in ambito sanitario (*Capitolo 2*), cui segue la descrizione della metodologia adottata per la conduzione della revisione della letteratura (*Capitolo 3*). I risultati emersi dall'analisi degli studi inclusi sono quindi presentati e organizzati secondo le dimensioni di stratificazione individuate (*Capitolo 4*). Tali evidenze vengono successivamente discusse in chiave critica, con approfondimento dei principali aspetti metodologici e dei gap ancora presenti nella letteratura scientifica (*Capitolo 5*). Infine, sono illustrati i limiti dello studio, le prospettive future di ricerca e le conclusioni (*Capitoli 6-7*).

CAPITOLO 2 - Contesto e quadro regolatorio

In questo capitolo si delineano le fondamenta concettuali e normative relative all'impiego dell'Intelligenza Artificiale (IA) in sanità, con particolare riferimento agli ambiti **ambulatoriale** e **chirurgico**. Nel sottocapitolo 2.1 vengono fornite le definizioni dei principali termini e approcci (IA, *Machine Learning*, *Deep Learning*, *Natural Language Processing*, *Reinforcement Learning*, *Clinical Decision Support Systems* e metodi di *Ricerca Operativa*), evidenziandone le differenze e le applicazioni chiave in sanità, in particolare per la pianificazione di attività ambulatoriali e chirurgiche. Il sottocapitolo 2.2 offre una panoramica delle **norme e delle guidance istituzionali** esistenti: si esaminano i documenti dell'Organizzazione Mondiale della Sanità (OMS) – ad esempio il rapporto *“Ethics and Governance of AI for Health”* (2021) e la pubblicazione *“Regulatory considerations on AI for health”* (2023) – quindi il quadro giuridico europeo con il recente **AI Act** dell'UE (approvato nel 2025), e infine la nuova **Legge italiana 23 settembre 2025 n. 132** sull'intelligenza artificiale. Per ciascuno di questi riferimenti si sintetizzano i principi e requisiti rilevanti per l'ambito sanitario (*con attenzione a temi come classificazione dei rischi, requisiti di trasparenza, sicurezza, auditabilità e supervisione umana*), discutendo anche le implicazioni operative per il personale sanitario che si trova a interagire con soluzioni di IA nei contesti assistenziali.

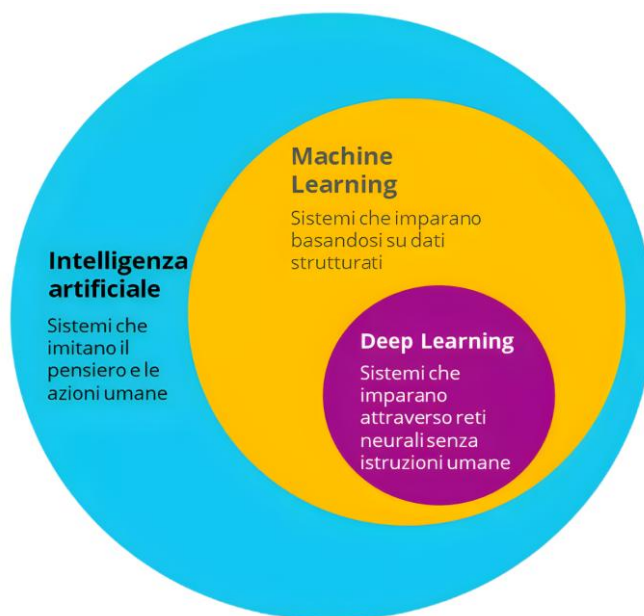


Figura 2 - Relazione tra Intelligenza Artificiale, Machine Learning e Deep Learning: l'IA rappresenta il campo più ampio, al cui interno il Machine Learning e, più specificamente, il Deep Learning costituiscono sottoinsiemi di sistemi di apprendimento automatico. Fonte: IONOS Digital Guide (2025).

2.1 Definizioni e concetti fondamentali

Intelligenza Artificiale (IA). Il termine *Intelligenza Artificiale* indica il campo dell'informatica e dell'ingegneria che sviluppa sistemi e algoritmi capaci di svolgere compiti tipicamente associati all'intelligenza umana, come l'analisi di informazioni, il ragionamento, l'apprendimento e il processo decisionale. Una definizione ampiamente condivisa descrive un sistema di IA come un sistema software progettato per **percepire l'ambiente, elaborare dati e produrre output (previsioni, raccomandazioni o decisioni)** in vista di obiettivi predefiniti (European Commission, 2024). In pratica, l'IA comprende una varietà di tecniche – *da algoritmi basati su regole e logica a sistemi che apprendono automaticamente dai dati* – con l'obiettivo di imitare aspetti dell'intelligenza umana o comunque di risolvere problemi complessi in modo *“intelligente”*. Nel settore sanitario, l'IA trova applicazione in numerosi ambiti: ad esempio nell'**elaborazione di diagnosi** e nel **riconoscimento di immagini mediche**, nel **supporto alle decisioni cliniche** (*clinical decision support system*), nella **sorveglianza epidemiologica**, fino al **miglioramento della gestione dei sistemi sanitari** (es. *ottimizzazione delle risorse, pianificazione delle cure*).

È importante sottolineare che l'IA non è una singola tecnologia, ma un *ombrello concettuale* sotto il quale ricadono diversi sottocampi specializzati, come di seguito definiti.

Machine Learning (ML) – Apprendimento automatico. Il *Machine Learning* è un sottoinsieme dell'IA che si focalizza sullo sviluppo di algoritmi in grado di **apprendere dai dati ed estrarre pattern o conoscenza** senza essere esplicitamente programmati per ogni decisione. In altre parole, un sistema ML *“impara”* da insiemi di dati storici a migliorare le proprie prestazioni su un compito specifico, affinando un modello interno che può poi generalizzare a nuovi dati. Esistono varie forme di ML: l'**apprendimento supervisionato** (*in cui il modello impara da esempi etichettati, ad es. immagini mediche con diagnosi note*), l'**apprendimento non supervisionato** (*in cui identifica autonomamente strutture o raggruppamenti nei dati, ad es. segmentare pazienti in cluster con bisogni simili*) e l'**apprendimento per rinforzo** (*di cui si dirà più avanti*). In sanità, il ML è già largamente impiegato: algoritmi di apprendimento automatico sono utilizzati per **predire eventi clinici o gestionali** (ad es. il rischio di complicanze o di riospedalizzazione di un paziente, la probabilità di

no-show mancata presentazione a un appuntamento, la durata prevista di un intervento chirurgico) e per **classificare grandi moli di dati clinici** (come nel caso di analisi di immagini radiologiche o di dati omici, ossia dati biologici ad alta dimensionalità relativi a genoma, trascrittoma o proteoma). La capacità del ML di trovare correlazioni complesse nei dati lo rende uno strumento potente per supportare l'ottimizzazione delle risorse: ad esempio, modelli predittivi possono anticipare quali giornate avranno un elevato carico di pazienti in un ambulatorio, o quali interventi chirurgici dureranno più del previsto, consentendo così una migliore pianificazione (World Health Organization, 2021). Va notato tuttavia che i modelli di ML, sebbene efficaci, funzionano come “**scatole nere**” (*in inglese black box*) in molti casi, rendendo talora difficile interpretarne le decisioni; per questo *trasparenza ed affidabilità* dei modelli sono criteri cruciali nella loro adozione in ambito clinico (si veda sottocapitolo 2.2) (World Health Organization, 2021).

Deep Learning (DL) – Apprendimento profondo. Il *Deep Learning* è una famiglia di tecniche di Machine Learning basata su **reti neurali artificiali con molti strati** (*deep neural networks*). Si parla di “apprendimento profondo” perché queste reti contengono strati gerarchici di neuroni artificiali che estraggono rappresentazioni di livello via via più alto dai dati grezzi. Il DL ha segnato progressi straordinari nell'ultimo decennio, trainando molte innovazioni dell'IA. In particolare, **reti neurali convolutive (CNN)** hanno raggiunto performance pari o superiori a quelle umane nel riconoscimento di immagini, trovando applicazione in diagnostica radiologica e istopatologia digitale (*es. rilevazione di noduli polmonari da radiografie, classificazione di biopsie*) (Zhou et al., 2024). **Reti neurali ricorrenti** e architetture derivate (come *LSTM* o *transformer*, queste ultime alla base dei moderni modelli linguistici) eccellono nell'analisi di dati sequenziali o testuali, aprendo la strada all'elaborazione di serie temporali cliniche (*ad es. trend di parametri vitali in terapia intensiva*) e all'elaborazione del linguaggio naturale in medicina. Grazie all'accesso a grandi quantità di dati (*big data*) e alla capacità computazionale odierna, il Deep Learning ha potuto alimentare sistemi di **diagnosi assistita** sempre più accurati e sistemi predittivi per la gestione sanitaria (*predictive analytics*). Ad esempio, in ambito chirurgico, modelli di DL sono impiegati per *predire la durata di un intervento* con maggiore precisione – analizzando dati storici, caratteristiche del

paziente e dell'équipe – così da ottimizzare il programma operatorio e ridurre i tempi morti delle sale operatorie (Getinge, 2024). Nonostante le potenzialità, il DL pone sfide relative alla **spiegabilità** (*in inglese Explainable*) (*spesso il perché di una certa decisione resta oscuro*) e alla **robustezza** (*sensibilità a dati di input leggermente diversi*): aspetti che la comunità scientifica e regolatoria stanno affrontando con linee guida per un'IA affidabile (si veda ancora sottocapitolo 2.2) (World Health Organization, 2021).

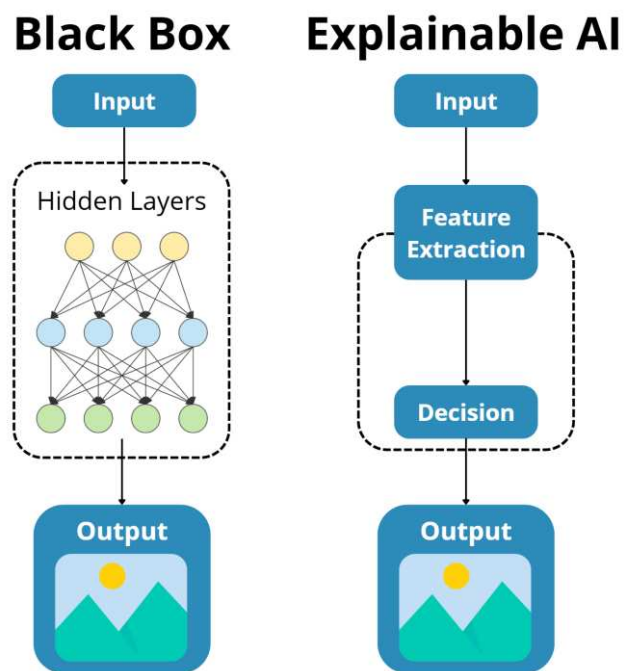


Figura 3 - Differenza concettuale tra modelli di tipo “Black Box” e “Explainable AI (XAI)”. Nei modelli di Deep Learning tradizionali, il processo di elaborazione interna dei dati attraverso gli strati nascosti non è interpretabile dall'essere umano (“black box”). I modelli di Intelligenza Artificiale spiegabile introducono invece fasi di estrazione e interpretazione delle caratteristiche, rendendo trasparente il processo decisionale.

Natural Language Processing (NLP) – Elaborazione del Linguaggio Naturale. Il *Natural Language Processing* è il campo dell'IA che si occupa dell'**interazione tra computer e linguaggio umano**, scritto o parlato. In ambito sanitario, le tecniche di NLP permettono di estrarre informazione utile dal vasto patrimonio di testi non strutturati che caratterizzano la pratica clinica: *cartelle cliniche elettroniche*, referti medici, appunti clinici, letteratura scientifica, questionari e persino conversazioni tra medico e paziente. Ad esempio, algoritmi di NLP possono analizzare automaticamente i referti di pronto soccorso per individuare i pazienti che necessitano di follow-up, identificare menzioni di sintomi o diagnosi in una cartella clinica per *popolare registri di patologia*, o supportare il **triage** valutando le

descrizioni dei sintomi fornite dai pazienti. Un caso particolare di NLP applicato alla gestione ambulatoriale sono i **chatbot medici** o sistemi di assistenza virtuale, che comprendono domande dei pazienti e forniscono risposte su procedure o preparazioni agli esami, contribuendo a snellire i processi informativi. Inoltre, grazie ai recenti avanzamenti nei modelli linguistici (es. reti *Transformer* come GPT), l’NLP consente di creare riassunti automatici di documenti clinici, di tradurre referti specialistici in un linguaggio più comprensibile per i pazienti, o di analizzare la *soddisfazione degli utenti* da survey testuali. Nei contesti di pianificazione e gestione, l’NLP può supportare l’**analisi della domanda sanitaria**: ad esempio, leggendo le richieste di prenotazione o di consulto per capire trend emergenti (*come un aumento di casi con certi sintomi che potrebbe anticipare un picco di domanda*). Pur essendo un campo distinto da ML e DL, l’NLP moderno fa ampio uso del Deep Learning (modelli di *language understanding*), combinando quindi le sfide di interpretabilità dei modelli con quelle di privacy e riservatezza dei dati testuali in sanità. Per questo, la governance dei progetti NLP clinici deve garantire che i dati sensibili vengano trattati in modo conforme alle normative (*si pensi al GDPR - General Data Protection Regulation - Regolamento UE 2016/679*) e che i modelli non introducano *bias* linguistici (*ad es. una minor accuratezza nel comprendere descrizioni di sintomi fornite in linguaggi colloquiali o da gruppi socio-culturali diversi*).

Reinforcement Learning (RL) – Apprendimento per Rinforzo. L’*apprendimento per rinforzo* è una particolare tecnica di ML in cui un “agente” artificiale impara a prendere decisioni **interagendo con un ambiente** e ricevendo un *feedback* sotto forma di ricompense o penalità. Invece di apprendere da un insieme fisso di dati etichettati, l’agente RL apprende gradualmente per tentativi ed errori, migliorando la propria politica d’azione per massimizzare la ricompensa cumulativa. Questo approccio è stato reso popolare dai successi nel gioco (es. *AlphaGo*), ma ha applicazioni promettenti anche in sanità, specialmente nei problemi di **decisione sequenziale complessa**. Ad esempio, in ambito terapeutico, algoritmi di RL sono stati sperimentati per ottimizzare regimi di dosaggio farmaceutico personalizzati, adattando le dosi in base alla risposta del singolo paziente nel tempo. Nel contesto della pianificazione sanitaria, il RL può essere impiegato per sviluppare sistemi di

supporto decisionale dinamico: si pensi alla gestione di una lista d'attesa chirurgica, dove a mano a mano che si liberano posti o che compaiono urgenze, un agente RL potrebbe proporre quale paziente chiamare o spostare, avendo come ricompensa una combinazione di obiettivi (*ridurre il tempo di attesa medio, massimizzare l'utilizzo delle sale, garantire equità nell'accesso*). Allo stesso modo, in un ambiente ambulatoriale, un sistema RL potrebbe apprendere una politica di *overbooking* ottimale (inserimento di pazienti extra in previsione di *no-show*) bilanciando il rischio di sovraffollamento con quello di slot sprecati. Un altro esempio è la **programmazione adattiva delle agende:** l'algoritmo può decidere di riservare alcuni slot per eventuali urgenze last-minute, o di allocare risorse (*es. personale di supporto, stanze*) in modo flessibile durante la giornata in base al flusso corrente di pazienti. Queste applicazioni sono ancora in fase sperimentale, ma mostrano il potenziale del RL nel "*domare*" la complessità operativa di sistemi sanitari soggetti a incertezza e variabilità. È importante, tuttavia, che tali sistemi RL in contesto clinico operino sotto supervisione e con adeguati vincoli di sicurezza: la **sperimentazione controllata** (*ad es. in sandboxes regolatorie, cfr. sottocapitolo 2.2.3*) è fondamentale prima che decisioni suggerite da un agente artificiale possano incidere sul percorso di cura di un paziente.

Clinical Decision Support Systems (CDSS) – Sistemi di supporto decisionale clinico.

I *Clinical Decision Support Systems* sono sistemi informatici progettati per assistere i professionisti sanitari nei processi decisionali clinici. Essi forniscono **conoscenza e informazioni intelligenti al punto di decisione**, tipicamente integrandosi con i flussi di lavoro clinici (*ad esempio all'interno del software della cartella clinica elettronica*). I CDS possono assumere forme diverse: da semplici *reminder* e avvisi (*come il richiamo automatico di linee guida o l>alert che segnala una potenziale interazione farmacologica pericolosa*) fino a complessi sistemi di raccomandazione diagnostica o terapeutica basati su algoritmi avanzati. Storicamente molti CDS erano basati su regole *if-then* o su alberi decisionali costruiti da esperti (*sistemi esperti*). Con l'avvento dell'IA, i CDS di nuova generazione incorporano sempre più spesso modelli di *Machine Learning*, capaci di analizzare dati clinici del paziente e suggerire decisioni personalizzate – ad esempio, un algoritmo che stima il rischio di deterioramento clinico di un paziente ospedalizzato e allerta in anticipo il team,

oppure un sistema che assiste nella scelta di una terapia antibiotica basandosi sul profilo del paziente e su database di letteratura. Nei contesti ambulatoriali e chirurgici, i CDS possono avere un impatto significativo: si pensi a un sistema di supporto che aiuta a **prioritizzare le liste d'attesa** in base sia all'urgenza clinica sia ad aspetti logistici, oppure un CDS che durante la visita ambulatoriale suggerisce esami integrativi al medico secondo le *best practice* (riducendo variabilità e ritardi diagnostici). In ambito chirurgico, un CDS può supportare il chirurgo nella fase decisionale pre-operatoria (*ad es. indicando, per un caso complesso, quali imaging o consulti specialistici aggiuntivi sono raccomandati prima di procedere*) o nel **decidere la priorità** tra diversi pazienti candidati alla sala operatoria quando compare uno slot libero. **Infermieri e ostetriche** spesso interagiscono con CDS, specialmente per la gestione di protocolli (*ad es. in ostetricia, sistemi che ricordano le tempistiche di monitoraggio per gravidanze a rischio, o che supportano il triage ostetrico*). Affinché i CDS siano efficaci e sicuri, è essenziale che siano **ben integrati nei sistemi informativi sanitari** e nel flusso di lavoro: l'informazione giusta deve essere presentata alla persona giusta, nel momento giusto. Inoltre, come sottolineano le normative recenti, i CDS – in quanto possibili *dispositivi medici software* – devono essere sviluppati e validati rigorosamente, e il personale sanitario deve essere formato al loro utilizzo critico (*cf. sottocapitolo 2.2 su requisiti di qualità, trasparenza e supervisione umana dei sistemi di IA in sanità*) (European Commission, 2024). Un CDS dovrebbe **supportare ma non mai sostituire** il giudizio clinico: questo principio, ribadito sia dall'OMS che dal legislatore, preserva il ruolo decisionale del professionista, il quale rimane il responsabile ultimo delle decisioni prese con l'ausilio del software (Gatto, 2025).

Metodi di Ricerca Operativa (Operations Research). Con *Ricerca Operativa (RO)* si intende l'insieme di metodi analitici – spesso matematici e computazionali – volti a **ottimizzare processi decisionali complessi**, tipicamente sotto vincoli di risorse limitate. La RO include tecniche come la **programmazione matematica** (lineare, intera, non lineare), gli **algoritmi di ottimizzazione combinatoria**, la **simulazione dei processi** e altre metodologie per risolvere problemi di allocazione ottimale di risorse, scheduling (*programmazione temporale di attività*) e progettazione di sistemi. In sanità, la Ricerca Operativa vanta una lunga tradizione di applicazioni:

dall'ottimizzazione dei turni del personale infermieristico, alla pianificazione delle reti di emergenza-urgenza (ad es. *posizionamento ottimale delle ambulanze sul territorio*), fino alla **programmazione delle sale operatorie** e delle **agende ambulatoriali**. Ad esempio, un classico problema di RO in ambito chirurgico è il **block scheduling**: come assegnare blocchi orari delle sale operatorie alle varie specialità o equipe, in modo da massimizzare l'utilizzo e minimizzare le cancellazioni, rispettando vincoli di disponibilità di chirurghi e anestesisti. Allo stesso modo, in ambito ambulatoriale, metodi di ottimizzazione possono aiutare a costruire agende bilanciate che riducano i tempi di attesa: modelli matematici possono determinare il mix ottimale di tipi di visite in un poliambulatorio (*prime visite vs controlli*), o la sequenza ottimale di pazienti tenendo conto di durata stimata e priorità clinica. Spesso la Ricerca Operativa fornisce soluzioni "esatte" o "ottimali" rispetto al modello matematico formulato; tuttavia, l'efficacia pratica dipende dalla bontà di tale modello e dei dati. Negli ultimi anni, si assiste a una convergenza tra IA e RO: da un lato, **algoritmi di apprendimento automatico** possono migliorare i parametri in ingresso ai modelli di ottimizzazione (ad es., *fornendo stime più accurate dei tempi chirurgici o delle probabilità di no-show, che poi un algoritmo di scheduling usa per decidere la sequenza ottimale*). Dall'altro lato, la RO tradizionale può essere impiegata per *ottimizzare l'uso di un modello di IA*: ad esempio, in un ospedale si potrebbe usare la programmazione lineare per decidere in quali reparti implementare prioritariamente una certa soluzione di IA, massimizzando il rapporto costi-benefici. Nei contesti di pianificazione operativa ambulatoriale e chirurgica, spesso si adottano **approcci ibridi IA + OR**, in cui componenti predittive (ML) e componenti prescrittive (ottimizzazione) lavorano in sinergia. Ad esempio, uno schema possibile è "*cluster-predict-optimize*": i pazienti in lista d'attesa vengono raggruppati in cluster omogenei (*magari tramite ML non supervisionato*), si predicono per ciascun cluster alcuni parametri (es. *tasso di no-show, durata media delle visite*) e infine un algoritmo di ottimizzazione determina la pianificazione migliore tenendo conto di queste informazioni. In conclusione, i metodi di Ricerca Operativa forniscono strumenti imprescindibili per affrontare il problema cronico delle liste d'attesa e del sovraccarico delle risorse sanitarie: integrati con l'IA, promettono soluzioni ancora più efficaci, ma richiedono attenta

validazione, modelli realistici e il coinvolgimento degli operatori sul campo per assicurare che i risultati siano attuabili e accettati.

Applicazioni in sanità, in particolare nella pianificazione ambulatoriale e chirurgica. Dopo aver introdotto i principali concetti, è utile riassumere **come queste tecnologie si applichino nei contesti sanitari di interesse.** Nella pianificazione **ambulatoriale**, l'obiettivo è ottimizzare l'utilizzo delle agende e delle risorse della clinica (*ambulatori, medici, infermieri*) per offrire ai pazienti tempi di attesa ridotti e un servizio efficiente. In questo ambito:

- **Machine Learning** vengono usati per *prevedere la domanda e supportare le decisioni di scheduling.* Ad esempio, modelli di ML supervisionato possono analizzare dati storici di prenotazioni per **predire i tassi di no-show** (*mancata presentazione*) per ogni giorno o per specifiche prestazioni. Conoscendo in anticipo che un certo giorno c'è alta probabilità di defezioni, il gestore dell'agenda può sovra-prenotare (*overbooking*) un numero adeguato di pazienti extra, mitigando l'impatto dei no-show senza aumentare troppo il rischio di code. Allo stesso modo, algoritmi predittivi aiutano a stimare quanti pazienti affluiranno in un dato periodo (*ad esempio, picchi influenzali che aumentano le visite non programmate*) così da allocare in anticipo più personale o ampliare le fasce orarie di visita.
- **Natural Language Processing** può contribuire ad esempio analizzando le *richieste di prenotazione online* o i *motivi di visita* riportati dai pazienti: attraverso l'NLP si possono categorizzare automaticamente le richieste (es. *"visita cardiologica" vs "controllo periodico"*), permettendo al sistema di dare priorità e assegnare la durata appropriata a ciascun appuntamento. Inoltre, chatbot basati su NLP possono interagire con i pazienti per raccogliere informazioni preliminari (*triage virtuale*) e indirizzarli al setting più adeguato (es. *visita ambulatoriale standard vs accesso diretto al PS*), contribuendo a *smistare* il flusso ed evitare intasamenti in agenda.
- **Reinforcement Learning** trova spazio in prototipi di sistemi di scheduling adattivo: ad esempio, un algoritmo RL potrebbe *imparare* la politica ottimale per gestire gli arrivi senza appuntamento (*walk-in*) in un poliambulatorio,

decidendo quando inserirli tra gli appuntamenti programmati in modo da minimizzare sia l'attesa dei walk-in sia il ritardo accumulato per le visite successive. Ricevendo come *reward* una funzione che penalizza le attese prolungate e i ritardi eccessivi, l'agente RL potrebbe affinare col tempo decisioni di inserimento sempre più efficienti. Si tratta di applicazioni sperimentali, ma concettualmente interessanti per situazioni in cui le regole fisse (es. *“vedere sempre un walk-in appena arriva”* oppure *“vederli tutti alla fine”*) non offrono la flessibilità necessaria a ottimi risultati.

- **Clinical Decision Support Systems** in ambito ambulatoriale possono aiutare a *prioritizzare gli appuntamenti* in base alla gravità. Ad esempio, un CDS potrebbe segnalare, tra i pazienti in lista per una visita specialistica, quelli con indicazioni cliniche che suggeriscono urgenza, assistendo i sanitari nel riservare slot ravvicinati a quei casi. Inoltre, i CDS possono suggerire *ottimizzazioni nella sequenza delle visite*: se integrati con modelli predittivi, potrebbero consigliare di mettere all'inizio della sessione pazienti che è più probabile non si presentino (*così da non scombinare l'agenda se saltano*) oppure di non collocare consecutivamente due visite potenzialmente lunghe (*per ridurre la probabilità di accumulo di ritardo*). Tali decisioni comunque richiedono convalida umana e buon senso clinico, rimarcando l'idea della IA come supporto e non sostituto.
- **Metodi di Ricerca Operativa** sono tradizionalmente impiegati per calcolare *la programmazione ottimale delle agende*. Ad esempio, un modello matematico di *Appointment Scheduling* può formulare il problema: *“dato un certo numero di appuntamenti di diverse tipologie da allocare in una settimana, con durate stimate e priorità note, come distribuire gli appuntamenti nei giorni/orari minimizzando l'attesa media e massimizzando l'utilizzo del tempo dei medici, rispettando i vincoli di orario di ciascun professionista e il limite di pazienti al giorno?”* Questi problemi si risolvono con algoritmi di ottimizzazione che forniscono soluzioni (ad es. attraverso *integer programming* o *heuristics*). Nella pratica, la Ricerca Operativa in ambito ambulatoriale produce *linee guida organizzative* (es. quanti appuntamenti di controllo vs prime visite pianificare al giorno per equilibrare domanda e offerta) e strumenti software che

suggeriscono slot ottimali per i nuovi appuntamenti inseriti, tenendo conto di criteri multipli (efficienza, equità, preferenze dei pazienti). Già oggi molte aziende sanitarie utilizzano moduli di ottimizzazione nei sistemi di prenotazione per proporre al paziente la data più vicina possibile che minimizza anche il carico di lavoro residuo.

Nel contesto **chirurgico**, la pianificazione si articola su più livelli: *programmazione strategica* (allocazione delle sale operatorie alle varie discipline o utilizzi, definizione dei turni), *programmazione tattica* (predisposizione del calendario operatorio, ad esempio la lista degli interventi per la prossima settimana) e *gestione giornaliera* (adattamenti in tempo reale alle variazioni, es. emergenze, ritardi, cancellazioni). Ciascuna di queste fasi può beneficiare di IA e metodi quantitativi:

- **Machine Learning/Deep Learning:** Un'area di impatto immediato è la *stima dei tempi chirurgici*. Modelli di apprendimento automatico addestrati su dati storici (*durata degli interventi passati con le relative caratteristiche*) sono in grado di prevedere con maggiore accuratezza la durata di un intervento programmato rispetto alle stime manuali dei chirurghi (Getinge, 2024). Variabili come il tipo di procedura, le condizioni specifiche del paziente, l'équipe coinvolta, l'ora del giorno, etc., possono essere considerate dal modello. Un miglioramento della stima da, ad esempio, $\pm 30\%$ di errore a $\pm 10\%$ può tradursi in un uso molto più efficiente delle sale: si riducono i *tempi morti* (sale vuote in attesa) e al contempo si riducono i *ritardi a fine giornata* che causano straordinari o cancellazioni dell'ultimo minuto (Capaccio, 2024). Altri modelli ML in ambito operatorio includono quelli per *prevedere le cancellazioni* (ad es. identificando pazienti a rischio di annullare l'intervento, magari per condizioni cliniche non stabilizzate o probabili problemi logistici – su ciò possono intervenire i coordinatori per tempo) e quelli per *predire complicanze post-operatorie* o necessità di terapia intensiva, informazioni che aiutano a pianificare posti letto e personale post-chirurgia.
- **Reinforcement Learning:** Può essere applicato alla *gestione dinamica delle liste operatorie*. Si immagina un agente che ogni giorno apprende quale intervento programmare successivamente in uno slot che si è liberato, basandosi sullo

stato corrente (lista d'attesa residua, priorità cliniche, tempi rimasti nella giornata, etc.) e su una funzione di ricompensa che premia l'utilizzo completo del tempo operatorio ma penalizza gli sforamenti oltre l'orario o l'attesa eccessiva di casi urgenti. Nel lungo termine, l'agente RL potrebbe identificare strategie (policy) migliori di quelle statiche – ad esempio decidendo **quale caso posticipare** se un intervento precedente si prolunga, in modo da minimizzare l'impatto sulla cura (evitando di posticipare un caso urgente o un paziente già preparato) e sull'efficienza. Questi sistemi, ancora sperimentali, mostrano la potenzialità dell'IA di adattarsi in tempo reale a situazioni variabili, che è un limite noto degli approcci deterministici classici.

- **Clinical Decision Support:** In sala operatoria, i CDS sono spesso integrati con monitoraggi e cartelle peri-operatorie per supportare decisioni cliniche (es. *suggerire protocolli di gestione del dolore, antibiotico profilassi, etc.*). Dal punto di vista organizzativo, un CDS può fornire al **manager del blocco operatorio** raccomandazioni su come aggiustare il programma quando ci sono imprevisti. Ad esempio, se salta un intervento all'ultimo momento, il sistema potrebbe suggerire quale paziente della lista d'attesa di pari priorità chiamare come rimpiazzo, tenendo conto di compatibilità di attrezzature e personale disponibile. Oppure, potrebbe segnalare che un certo chirurgo finirà in anticipo la lista della sua sala e proporre di spostare un caso da un'altra sala in ritardo, ottimizzando l'uso di tutte le sale (ovviamente previa valutazione del personale coinvolto). Tali funzioni di supporto decisionale **multifattoriale** sono possibili combinando modelli predittivi (*che forniscono stime aggiornate dei tempi*) con regole di business e vincoli clinici codificati.
- **Ricerca Operativa:** È il pilastro tradizionale della programmazione chirurgica. Algoritmi di *scheduling* producono soluzioni per l'**assegnazione ottimale delle sale** alle specialità e la **sequenza degli interventi**. Ad esempio, un classico problema risolto con RO è: data una serie di interventi da pianificare in una settimana, con durate stimate e priorità (*ed eventualmente urgenze che potrebbero inserirsi*), assegnare ciascun intervento a una sala e a un giorno minimizzando la somma dei tempi di attesa ponderati per la priorità, e

garantendo che il carico di ore non ecceda il monte ore disponibile di ciascuna sala per giorno. Questo può essere modellato come un problema di ottimizzazione intera e risolto con solutori avanzati; in contesti molto complessi si ricorre a metodi euristici o meta-euristici (es. *genetic algorithms*, *simulated annealing*) che trovano buone soluzioni in tempi ragionevoli. La RO viene anche utilizzata per definire **piani operativi robusti**: ad esempio, includendo margini di tempo cuscinetto per assorbire variazioni, o risolvendo problemi di ottimizzazione **stocastica** che tengano conto dell'incertezza nelle durate (generando piani validi su vari scenari). Tutto ciò mira a ridurre le **liste d'attesa chirurgiche** e a utilizzare al meglio le costose risorse operatorie. Studi riportano che l'introduzione di sistemi di ottimizzazione per il calendario operatorio può incrementare significativamente il numero di interventi completati settimanalmente senza aumentare le risorse, semplicemente riducendo tempi morti e inefficienze (Capaccio, 2024). Va però evidenziato che una *soluzione "ottima" sulla carta* deve poi essere attuabile nella realtà: per questo i modelli devono includere vincoli realisti (*pause del personale, preferenze dei chirurghi, requisiti di sanificazione sale, etc.*) e soprattutto la pianificazione deve essere **condivisa e accettata** dal team chirurgico. Il coinvolgimento attivo di **infermieri di sala, coordinatori chirurgici e altri operatori** nella definizione dei criteri di ottimizzazione e nella revisione dei piani proposti è fondamentale per il successo dell'implementazione (questo punto sarà ripreso nelle implicazioni per la pratica infermieristica, sottocapitolo 2.2.3).

In sintesi, IA, ML/DL, NLP, RL, CDS e metodi di Ricerca Operativa offrono ciascuno contributi specifici e spesso complementari al miglioramento dell'efficienza e dell'efficacia in ambito sanitario. Nel contesto di questa tesi, l'attenzione è rivolta all'uso di tali soluzioni per **ottimizzare le risorse ambulatoriali e chirurgiche** e valutarne l'impatto sugli esiti. Le tecnologie descritte non vanno considerate in isolamento: le applicazioni più efficaci tendono ad integrarle (es. *predictive analytics* + *ottimizzazione matematica* + *decision support*). Tuttavia, come approfondiremo nel prossimo capitolo, la letteratura evidenzia risultati eterogenei sull'"**efficacia**" di queste soluzioni; da qui l'importanza di comprendere bene il **contesto normativo** e

regolatorio in cui si inseriscono, che può facilitare o ostacolare l'adozione e che pone paletti imprescindibili (etici e legali) al loro utilizzo.

2.2 Norme e guidance su IA in ambito sanitario

L'uso dell'IA in sanità solleva questioni cruciali di **etica, sicurezza, responsabilità e tutela dei diritti**, il che ha spinto organismi internazionali e autorità a emanare linee guida e normative specifiche. In questa sezione si esaminano: (i) le principali **guidance internazionali**, con particolare attenzione al ruolo dell'Organizzazione Mondiale della Sanità (OMS) che nel 2021 ha pubblicato le prime linee guida etiche globali sull'IA in salute e nel 2023 ha elencato considerazioni chiave per la regolamentazione dell'IA in ambito sanitario; (ii) il quadro normativo **europeo**, incentrato sul **Regolamento UE sull'Intelligenza Artificiale (AI Act)** approvato in forma consolidata nel 2025, che introduce un approccio basato sul rischio e requisiti stringenti per sistemi di IA, con particolare impatto sugli utilizzi in sanità; (iii) il recente **quadro normativo italiano**, ovvero la Legge 23 settembre 2025 n. 132, prima legge nazionale organica sull'IA, che dedica articoli specifici all'impiego in ambito sanitario e alla protezione dei pazienti e operatori. Si metteranno in luce i **principi comuni** (es. *centralità della supervisione umana, trasparenza, equità algoritmica*) e le **differenze** tra queste fonti, nonché le implicazioni operative per i professionisti sanitari, che dovranno lavorare con sistemi di IA rispettando tali regole.

2.2.1 Guidance dell'Organizzazione Mondiale della Sanità: etica e regolamentazione

Principi etici dell'OMS per l'IA in salute (rapporto 2021). Nel giugno 2021 l'OMS ha pubblicato il suo primo rapporto globale su etica e governance dell'Intelligenza Artificiale in ambito sanitario, intitolato *"Ethics and governance of artificial intelligence for health"*. Questo documento – frutto di due anni di consultazioni con esperti internazionali – ha sottolineato da un lato le enormi opportunità dell'IA per migliorare l'accesso e la qualità delle cure, dall'altro i rischi e le sfide che l'IA comporta se non governata adeguatamente (World Health Organization, 2021).

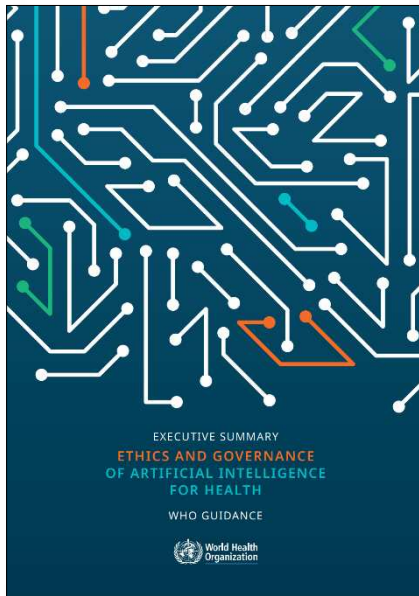


Figura 4 - Copertina del documento "Ethics and Governance of Artificial Intelligence for Health" dell'Organizzazione Mondiale della Sanità (WHO, 2021), che fornisce linee guida globali per un utilizzo etico, sicuro e trasparente dell'intelligenza artificiale in ambito sanitario.

Il rapporto identifica vari scenari positivi: *maggior velocità e accuratezza diagnostica, terapie più personalizzate, ricerca biomedica accelerata, sistemi sanitari più efficienti*, e anche *empowerment* dei pazienti nel gestire la propria salute (World Health Organization, 2021). Parallelamente, vengono evidenziate criticità quali i possibili **bias algoritmici** che possono

condurre a decisioni inique, l'utilizzo non etico dei dati sanitari (*con potenziali violazioni della privacy*), i rischi per la sicurezza del paziente derivanti da errori dei sistemi di IA, nonché il pericolo che un'adozione non adeguatamente governata dell'IA possa sottrarre risorse a interventi di sanità pubblica consolidati o amplificare disuguaglianze, ad esempio qualora i sistemi risultino performanti solo per popolazioni simili a quelle su cui sono stati addestrati. Per massimizzare benefici e minimizzare rischi, l'OMS propone **sei principi guida** che dovrebbero informare la progettazione e l'uso dell'IA in sanità in tutti i Paesi (World Health Organization, 2021):

- 1. Protezione dell'autonomia umana** – Le decisioni sanitarie devono **restare sotto controllo umano**: l'IA non deve sostituire la discrezionalità e il giudizio del personale sanitario. Inoltre, vanno tutelati la privacy e il diritto al consenso informato dei pazienti riguardo all'uso dei loro dati e all'intervento di sistemi di IA nel loro percorso di cura. In pratica, ciò implica informare il paziente se un certo esito (*diagnosi o terapia*) è stato coadiuvato da un algoritmo, e assicurarsi che il paziente (o il medico) possa opporsi o chiedere alternative. Questo principio difende la dignità e l'agenzia dell'individuo, evitando che l'IA riduca i pazienti a oggetti passivi di decisioni automatizzate.
- 2. Promozione del benessere, sicurezza e interesse pubblico** – Lo sviluppo di sistemi di IA per la salute deve mirare primariamente al **bene del paziente e della collettività**. I progettisti devono garantire che tali sistemi rispettino

standard di **sicurezza, accuratezza ed efficacia** appropriati per gli specifici impieghi clinici. Ciò comporta ad esempio che un algoritmo diagnostico sia validato clinicamente con rigore paragonabile a quello richiesto per un nuovo dispositivo medico o farmaco. Inoltre, devono esistere misure di controllo di qualità e miglioramento continuo nell'uso dell'IA, così che eventuali errori o cali di performance vengano rapidamente identificati e corretti. In una parola, l'IA in sanità deve *fare bene e non fare male*, in coerenza col principio ippocratico, e il suo impiego va giustificato dal valore aggiunto in termini di esiti di salute.

3. Trasparenza, spiegabilità e intelligibilità – Secondo l'OMS, è fondamentale che vi sia **trasparenza** sia nella fase di progettazione sia nell'implementazione dei sistemi di IA sanitari. Questo significa che informazioni adeguate sul funzionamento dell'algoritmo, sul tipo di dati utilizzati per addestrarlo e sul suo scopo devono essere pubblicamente disponibili o comunque accessibili alle autorità e agli utenti (*professionisti e pazienti*) in forma comprensibile. La **spiegabilità** e **intelligibilità** implicano che per quanto possibile i modelli di IA dovrebbero fornire motivazioni comprensibili per le loro raccomandazioni (*ad esempio evidenziando quali fattori del paziente hanno pesato in un certo punteggio di rischio*). La trasparenza si estende anche al coinvolgimento degli stakeholder: l'OMS auspica che vi sia consultazione pubblica e dibattito su come e dove certe tecnologie andrebbero usate. Questo principio di trasparenza prefigura requisiti che saranno poi recepiti in normative come l'AI Act europeo (*che infatti richiede documentazione tecnica e informative per gli utenti dei sistemi di IA ad alto rischio, cfr. sottocapitolo 2.2.2*).

4. Responsabilità e accountability – Nonostante l'IA esegua compiti specifici, la **responsabilità ultima** per le decisioni deve rimanere in capo a sviluppatori, fornitori e professionisti coinvolti. L'OMS sottolinea che devono essere predisposti meccanismi efficaci perché individui o gruppi danneggiati da decisioni basate su algoritmi possano porre domande, contestare e ottenere rimedi (*redress*). Questo implica definire chiaramente chi risponde in caso di errore: ad esempio, se un software di supporto diagnostico fallisce una diagnosi evidente causando danno, il paziente deve poter ottenere giustizia – e ciò

richiede norme sulla **responsabilità e professionalità** che non lascino vuoti. Allo stesso tempo, sviluppatori e utilizzatori devono adottare una *cultura della responsabilità*: validare i sistemi, monitorarne le performance, segnalare incidenti (*questo sarà poi un obbligo in base al Regolamento UE*). Un importante corollario è la **formazione appropriata**: chi usa l'IA (*medici, infermieri, ecc.*) deve essere addestrato a farlo correttamente, e chi la sviluppa deve includere expertise clinica, così che nessuno possa deresponsabilizzarsi scaricando colpe "*sul computer*". In sintesi, l'IA non esenta dal dovere di diligenza i professionisti; al contrario, l'uso di IA richiede un *surplus* di responsabilizzazione, con nuove competenze e protocolli di controllo.

5. **Inclusività ed equità** – L'IA per la salute deve essere progettata in modo da promuovere la più ampia **inclusione** possibile e un accesso equo ai suoi benefici. Ciò significa evitare che sistemi di IA funzionino bene solo per determinate popolazioni (*ad es. pazienti di Paesi ricchi*) e male per altre: i dataset di training dovrebbero riflettere la diversità in termini di età, genere, etnia, status socio-economico, ecc., per prevenire *bias* discriminatori. Inoltre, l'innovazione in ambito IA deve essere indirizzata anche ai bisogni dei contesti meno avvantaggiati: l'OMS richiama esplicitamente la necessità che **i Paesi a basso e medio reddito beneficino dell'IA** e non rimangano indietro (*ad esempio, adattando tecnologie alle risorse limitate, o prevedendo trasferimenti di conoscenza*). Inclusività significa anche **accessibilità**: sistemi di IA in sanità dovrebbero idealmente essere utilizzabili indipendentemente da disabilità (*es. interfacce compatibili con tecnologie assistive per operatori o pazienti con disabilità visive/uditive*), e considerare differenze linguistiche e culturali. In ambito operativo, il principio di equità implica ad esempio che un algoritmo per la gestione delle liste d'attesa **non privilegi né penalizzi sistematicamente** un gruppo di pazienti sulla base di attributi irrilevanti (*p.e. lo status socioeconomico*) – eventuali disparità vanno rilevate e corrette.
6. **AI responsabile e sostenibile** – L'ultimo principio afferma che i sistemi di IA dovrebbero essere **responsive**, cioè adattabili in base ai *feedback* durante l'uso, e **sostenibili** nel tempo. *Responsiveness* significa monitorare continuamente le

prestazioni del sistema in condizioni reali, per verificare che soddisfi le aspettative e requisiti, apportando aggiustamenti se emergono problemi o nuovi bisogni. Ad esempio, se un modello predittivo perde accuratezza perché cambiano i dati (concetto di *data drift*), occorre provvedere ad aggiornarlo. La *sostenibilità* ha vari significati qui: uno riguarda l'impatto ambientale – l'OMS invita a progettare IA **efficienti dal punto di vista energetico**, dato che i modelli di Deep Learning possono avere un'impronta carbonica elevata. Un altro aspetto è la **sostenibilità sociale**: governi e aziende dovrebbero prevedere misure per gestire l'impatto dell'automazione sul lavoro, ad esempio preparando **percorsi di riqualificazione per i lavoratori sanitari** i cui compiti vengono automatizzati, e prevenendo effetti negativi sull'occupazione. Nel contesto infermieristico, questo significa che se certe attività di routine (es. *monitoraggio parametri, trascrizione di appunti*) verranno svolte da IA, gli infermieri vanno formati per svolgere attività a più alto valore aggiunto e comunque per sorvegliare le macchine. In definitiva, l'OMS lega la **sostenibilità** all'idea di **governance adattiva**: l'ecosistema AI-sanità deve potersi evolvere con regole e strategie che bilancino innovazione e tutela dei valori umani nel lungo termine.

Questi sei principi dell'OMS costituiscono un quadro etico di riferimento. Essi sono volutamente generali, destinati a guidare sia i decisori politici che i progettisti e gli utilizzatori di IA. Li ritroveremo in gran parte riflessi nelle normative "*dure*" come il Regolamento UE e la legge italiana. Vale la pena notare che l'OMS, pur enfatizzando i benefici, *mette in guardia contro facili entusiasmi*: raccomanda ai Paesi di non trascurare gli investimenti di base per la copertura sanitaria universale inseguendo soluzioni IA di moda. Questo monito è particolarmente rilevante per chi – manager sanitari o politici – vede nell'IA una soluzione rapida al problema annoso delle liste d'attesa o carenza di personale: le evidenze vanno interpretate con rigore e le soluzioni implementate in modo complementare, non sostitutivo, al potenziamento strutturale dei servizi sanitari.

Considerazioni regolatorie OMS sull'IA in salute (rapporto 2023). Proseguendo il suo impegno, nell'ottobre 2023 l'OMS ha rilasciato una nuova pubblicazione

intitolata “Regulatory considerations on AI for health”, elaborata da un gruppo di lavoro internazionale sulla regolamentazione dell’IA in sanità (Baker McKenzie, 2023).

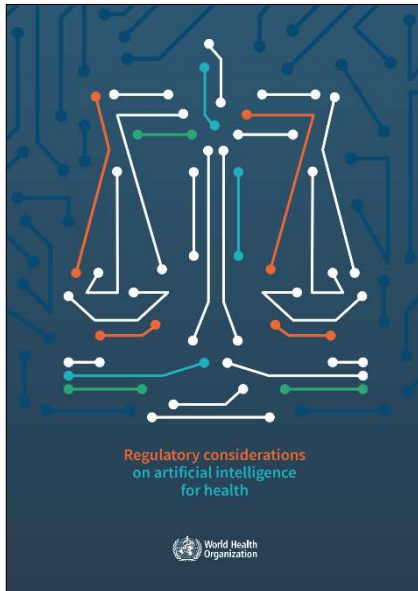


Figura 5 - Copertina del documento “Regulatory considerations on Artificial Intelligence for Health” pubblicato dall’Organizzazione Mondiale della Sanità (WHO, 2023), che delinea i principi e le raccomandazioni per la regolamentazione dei sistemi di intelligenza.

Questo documento si propone di fornire raccomandazioni pratiche ai *regolatori* e agli **stakeholder** (*sviluppatori, produttori, professionisti sanitari, pazienti*) su come affrontare le sfide specifiche poste dall’IA nell’ambito dei dispositivi medici, dei software clinici e delle applicazioni sanitarie. Vengono individuati **sei aspetti chiave** su cui costruire

framework regolatori solidi, ciascuno corredato da raccomandazioni operative (World Health Organization, 2023):

- 1. Documentazione e trasparenza** – È essenziale mantenere una documentazione accurata e completa sullo sviluppo e la validazione dei sistemi di IA in sanità. Ogni algoritmo dovrebbe avere **dichiarato l’“intended use”** (scopo clinico previsto) e traccia del processo di sviluppo, in modo da poter ricostruire i passi seguiti e verificarne la correttezza. La raccomandazione è di seguire un *approccio graduale*: maggiore è il rischio associato al sistema, più dettagliata e rigorosa dev’essere la documentazione richiesta (*principio proporzionale al rischio*). Questa enfasi sulla trasparenza tecnica mira a instaurare fiducia e a **facilitare la valutazione regolatoria** dei sistemi (*ad es. da parte di enti notificati, agenzie del farmaco o dispositivi, ecc.*). In pratica, significa che gli sviluppatori dovrebbero pre-specificare e documentare ogni fase: dalla selezione dei dataset di addestramento, all’architettura del modello, ai risultati dei test di validazione, rendendo tali informazioni disponibili alle autorità competenti.
- 2. Gestione del rischio e ciclo di vita** – I rischi associati ai sistemi di IA (*come vulnerabilità di cybersecurity, bias algoritmici o possibili malfunzionamenti*

clinici) devono essere considerati e affrontati **lungo tutto il ciclo di vita del prodotto**, dalla progettazione iniziale fino alla dismissione. Ciò richiede l'implementazione di un **approccio olistico di risk management**, integrato nel processo di sviluppo (**by design**) e nell'utilizzo operativo. L'OMS raccomanda di adottare le pratiche tipiche dei dispositivi medici: analisi dei rischi, mitigazioni, verifiche e riesami periodici, tenendo conto del contesto completo d'uso del sistema. In particolare, andrebbero affrontati rischi di **cybersecurity** e **vulnerabilità** (*un sistema di IA potrebbe essere attaccato o manipolato, con impatto diretto su diagnosi/terapie*), rischi di **errore o sotto-performance** (*es. algoritmi che degradano quando applicati a pazienti leggermente diversi da quelli di training*) e **bias** (*errori sistematici verso certi gruppi*). La cultura della **"safety by design"** qui è cruciale: integrare sin dall'inizio la sicurezza, anziché cercare di aggiustarla a prodotto finito. Per i sistemi di IA ad alto rischio, l'OMS suggerisce inoltre una fase di monitoraggio intensivo post-marketing, per captare tempestivamente eventi avversi o segnali di rischio inattesi.

- 3. Uso previsto e validazione** – La regolamentazione deve assicurare che di ogni sistema di IA siano chiari e pubblici **lo scopo previsto e le evidenze di validazione clinica** a supporto. Gli sviluppatori dovrebbero fornire **informazioni trasparenti sul dataset di training**, includendo composizione e caratteristiche dei dati usati per addestrare l'algoritmo. Ciò consente di valutare se, ad esempio, un modello per predire complicanze chirurgiche è stato addestrato su una popolazione comparabile a quella dove si intende usarlo (*evitando di applicarlo a contesti totalmente differenti*). Inoltre, la performance va dimostrata non solo sui dati di training ma anche su **dataset esterni di validazione** – idealmente provenienti da ambienti e popolazioni diverse – per provare la **generalizzabilità**. Le raccomandazioni prevedono requisiti di validazione **proporzionati al rischio**: per sistemi a basso rischio potrebbero bastare studi più snelli, mentre per quelli ad alto impatto (*es. un sistema che supporta decisioni diagnostiche critiche*) servono evidenze robuste, potenzialmente anche trial clinici comparativi. Un'altra best practice indicata è prevedere un periodo di **sorveglianza attiva post-distribuzione** (*post-market surveillance*) soprattutto per gli AI ad alto rischio, in cui si monitora l'algoritmo

nella pratica reale e si raccolgono feedback per eventuali aggiustamenti. Questo anticipa i requisiti poi codificati nel regolamento UE, che infatti impone ai fornitori di IA ad alto rischio di istituire sistemi di monitoraggio post-vendita e reportistica di incidenti.

4. Qualità dei dati – “I dati sono la linfa vitale dell’IA”: l’OMS sottolinea che la **qualità dei dati di input** influisce direttamente su sicurezza ed efficacia dei sistemi. Pertanto, i regolatori e sviluppatori devono accertarsi che **i dataset utilizzati siano di qualità sufficientemente alta per lo scopo previsto**. Ciò include aspetti come accuratezza, completezza, rappresentatività e aggiornamento dei dati clinici. Uno scenario comune in sanità è dover usare dati raccolti per scopi clinici (*es. cartelle cliniche elettroniche*) che possono contenere errori o inconsistenze: un modello addestrato su dati rumorosi avrà performance scadenti. Le raccomandazioni invitano a svolgere *valutazioni rigorose pre-release* dei sistemi di IA per assicurarsi che **non amplifichino bias o errori presenti nei dati**. Ad esempio, testare un algoritmo di triage su simulazioni per verificare che non discrimini pazienti anziani o di un certo sesso, anche se nel training set vi fosse una distribuzione non bilanciata. Se emergono problemi di qualità dati, occorre intervenire con *data cleaning*, raccolta di dati addizionali o tecniche per mitigare i bias. Viene incoraggiata anche la **creazione di ecosistemi collaborativi** per la condivisione di dataset di buona qualità – *nel rispetto delle normative privacy* – così da evitare che ogni attore si basi su piccoli dataset locali poco rappresentativi. In sintesi, data quality by design e testing anti-bias sono pilastri che un regolatore dovrebbe esigere prima di autorizzare un sistema di IA per l’uso clinico.

5. Privacy e protezione dei dati – Il rispetto della privacy e la protezione dei dati sanitari deve essere considerato **fin dall’inizio** nella progettazione, sviluppo e implementazione di sistemi di IA. I dati sanitari sono **dati personali sensibili** e come tali soggetti a normative stringenti (*si pensi al GDPR in Europa*); dunque chi sviluppa IA deve conoscere a fondo e **conformarsi al quadro legale vigente in materia di dati personali**. Le raccomandazioni OMS includono: incorporare misure di **privacy by design** (*ad es. minimizzazione dei dati, uso di tecniche di*

anonimizzazione/pseudonimizzazione, crittografia, controllo degli accessi) e condurre valutazioni d'impatto sulla protezione dati (DPIA - *Data Protection Impact Assessment*) per nuovi sistemi, specie se trattano big data sanitari. Inoltre, la *cybersecurity* viene qui richiamata: proteggere i dati dall'accesso non autorizzato e prevenire attacchi è parte integrante della compliance. L'OMS va oltre la legge e menziona anche **considerazioni etiche ulteriori**: ad esempio chiedersi se l'uso previsto dei dati è effettivamente spiegato e accettato dai pazienti, e garantire che vi sia trasparenza su eventuali utilizzi secondari dei dati. In definitiva, l'IA sanitaria deve guadagnarsi la fiducia del pubblico mostrando il massimo rigore nel trattare le informazioni personali – un punto su cui insiste anche la legge italiana 132/2025, che nel suo Art.9 richiama l'aderenza ai migliori standard di protezione dei dati sensibili in sanità (*cf. sottocapitolo 2.2.3*).

- 6. Coinvolgimento e collaborazione tra stakeholder** – L'ultimo asse evidenziato è che per garantire un'IA sanitaria conforme e *di qualità* è necessaria un'ampia **collaborazione tra tutti gli attori chiave**. Ciò significa coinvolgere attivamente **operatori sanitari, pazienti, sviluppatori, industria, enti regolatori e accademia** in un dialogo continuo. Le raccomandazioni incoraggiano la creazione di piattaforme di confronto, tavoli tecnici e "**sandbox**" regolatorie dove innovatori e regolatori possano sperimentare insieme soluzioni e linee guida. Ad esempio, sviluppare *hub* o *reti* in cui ospedali, università e aziende condividono esperienze sull'implementazione di IA in reparti pilota, informando le autorità su cosa funziona e cosa no in pratica, e ricevendo indicazioni su come rispettare i requisiti normativi. Questo approccio cooperativo serve a **snellire la supervisione** – rendendo il processo di conformità più efficiente – e ad assicurare che gli strumenti sviluppati rispondano veramente ai bisogni di chi li userà (*clinici e pazienti*). Un esempio potrebbe essere l'elaborazione congiunta di linee guida tecniche su come validare un algoritmo predittivo di rischio clinico: se al tavolo siedono medici, infermieri, statistici, enti notificati e aziende, sarà più facile concordare metriche e protocolli accettabili a tutti. Questa collaborazione contribuisce anche all'**armonizzazione** internazionale: l'OMS suggerisce di guardare alle esperienze di vari Paesi e di lavorare verso standard

globali (*ad esempio standard ISO specifici per IA medica, ecc.*), per evitare frammentazione e duplicazione degli sforzi.

Complessivamente, le **considerazioni regolatorie OMS del 2023** integrano i principi etici del 2021 con una prospettiva molto pratica su *come costruire fiducia e sicurezza* attorno all'IA in sanità. Esse anticipano molti aspetti che si ritrovano nell'AI Act europeo (*risk management, documentazione tecnica, data quality, supervisione umana*) e nella legge italiana (*trasparenza verso pazienti, sicurezza by design, formazione del personale*). Il messaggio chiave è che **una buona governance dell'IA sanitaria richiede sia regole chiare sia collaborazione multi-stakeholder**. Per i professionisti sanitari – tra i principali utilizzatori e gestori operativi delle tecnologie in corsia – tali raccomandazioni implicano, ad esempio, che dovranno essere coinvolti nella definizione di requisiti funzionali dei sistemi (*per assicurare usabilità*), che dovranno segnalare attivamente eventuali malfunzionamenti o errori algoritmici osservati nella pratica (*parte della post-market surveillance*), e che dovranno acquisire familiarità con nozioni base di gestione del rischio informatico e di protezione dati (*per rispettare protocolli di sicurezza nell'uso quotidiano di strumenti di IA*).

2.2.2 Il quadro normativo europeo: l'AI Act (Regolamento UE 2024/1689) e il settore sanitario



A livello dell'Unione Europea, la crescente importanza dell'Intelligenza Artificiale ha portato alla definizione di un **quadro normativo organico e vincolante**, culminato nell'approvazione del cosiddetto **AI Act (Regolamento UE 2024/1689, recante norme armonizzate sull'Intelligenza Artificiale)**, che rappresenta il **primo quadro legislativo completo al mondo sull'IA**, ed è finalizzato a garantire che i sistemi di IA immessi sul mercato europeo siano **sicuri, affidabili e rispettosi dei diritti fondamentali** (*European Commission, 2024*).

Dopo negoziati iniziati con la proposta della Commissione nel 2021, il testo è

stato consolidato e approvato tra la fine del 2023 e l'inizio del 2024, con entrata in vigore il 1° agosto 2024; essendo un regolamento, diventerà direttamente applicabile in tutti gli Stati membri dopo un periodo transitorio (24-36 mesi a seconda delle disposizioni) (European Commission, 2024). Nel 2025 le istituzioni e gli operatori si trovano quindi nella fase di preparazione all'applicazione pratica di queste norme, che saranno pienamente efficaci dal 2026 in poi.

Il **principio cardine dell'AI Act** è un **approccio basato sul rischio** (*risk-based approach*). In concreto, il Regolamento **classifica i sistemi di IA in quattro categorie di rischio** e prevede obblighi proporzionati al livello di rischio identificato:

- **Rischio inaccettabile (proibito)**: include quei sistemi di IA che per loro natura costituiscono una minaccia chiara a sicurezza, diritti e libertà delle persone. L'AI Act **vieta esplicitamente** una serie di pratiche di IA, tra cui: sistemi che manipolano il comportamento umano tramite tecniche subliminali causando danni, sistemi che sfruttano vulnerabilità di gruppi deboli (*minori, persone fragili*) per influenzarli in modo dannoso, sistemi di **social scoring** (*punteggi sociali stile "credito sociale"*), alcuni sistemi di sorveglianza biometrica di massa in tempo reale, ecc. Sebbene queste proibizioni riguardino principalmente ambiti come ordine pubblico e sorveglianza, ve ne sono alcune rilevanti per la sanità: ad esempio, sarebbe proibito un sistema che **valuti l'ammissibilità di un individuo a cure o assicurazioni sanitarie su basi arbitrarie o discriminatorie** (*ricadrebbe nel concetto di social scoring se penalizza persone per caratteristiche personali non correlate alla salute*). In generale comunque, le applicazioni tipiche sanitarie (*diagnosi, supporto clinico, gestione ospedaliera*) non rientrano nelle pratiche bandite, a meno di usi estremi (*ad es. un ipotetico sistema che negasse automaticamente l'accesso a terapie costose a pazienti con basso "valore sociale": scenario eticamente inaccettabile e proibito sia dal diritto alla salute sia dall'AI Act*).
- **Rischio Alto (regolamentato stringente)**: è il fulcro del Regolamento. Sistemi di IA **"ad alto rischio"** sono quelli che possono avere **impatti significativi su salute, sicurezza o diritti fondamentali** degli individui. L'AI Act definisce due grandi categorie di IA ad alto rischio: (1) sistemi di IA che sono **componenti di**

sicurezza di prodotti regolati da specifiche normative UE (*elencati nell'Allegato II del Regolamento*) – in questa categoria rientrano, **per la sanità, i dispositivi medici** e dispositivi medico-diagnostici in vitro che incorporano IA (*esempi: software diagnostici classificati come dispositivo medico, robot chirurgici con funzioni AI*); (2) sistemi di IA impiegati in **ambiti critici** (*Allegato III – AI ACT*), tra cui diverse voci attinenti alla sfera sanitaria e sociale. L'Allegato III infatti include: *“IA per l’assegnazione di servizi essenziali pubblici e privati (ad es. sistemi usati da autorità pubbliche per valutare l’ammissibilità a prestazioni sociali, o per il triage delle chiamate di emergenza sanitaria)”*, *“IA come componente di sicurezza di infrastrutture critiche (ad es. applicazioni in chirurgia robotica)”*, e altre voci per settori come istruzione, occupazione, forze dell’ordine, migrazione, giustizia. Notiamo quindi che **molte applicazioni sanitarie dell’IA ricadono automaticamente nel regime di alto rischio**: se il software è un dispositivo medico secondo MDR/IVDR, è considerato high-risk AI; se il sistema serve a smistare pazienti per l’accesso a cure urgenti o allocate risorse sanitarie scarse, è high-risk come *“accesso a servizi essenziali”*. Ad esempio, un algoritmo che prioritizza le liste d’attesa per interventi potrebbe essere considerato ad alto rischio se usato in ambito pubblico, perché influenza l’accesso effettivo alle cure (*che è un servizio essenziale*). Anche i sistemi di supporto alle diagnosi e terapie, essendo dispositivi medici software (*Software as a Medical Device- SaMD*), ricadono nel campo ad alto rischio e dovranno soddisfare i requisiti imposti dall’AI Act oltre a quelli della regolamentazione medica già esistente.

- **Rischio limitato (obblighi di trasparenza)**: include sistemi che interagiscono direttamente con gli esseri umani o generano contenuti potenzialmente percepiti come reali. L’AI Act impone in questo caso obblighi di trasparenza, affinché l’utente sia consapevole di interagire con un sistema di IA e non con un essere umano. Ciò riguarda ad esempio chatbot, assistenti virtuali sanitari o strumenti di generazione automatica di referti e immagini cliniche. Pur non essendo sottoposti a valutazione di conformità, questi sistemi devono comunque garantire chiarezza e correttezza nell’informazione, evitando il rischio di confusione o affidamento eccessivo da parte dell’utente.

- **Rischio minimo (nessun obbligo specifico):** comprende la grande maggioranza delle applicazioni di IA attualmente in uso, considerate prive di impatti significativi sui diritti o sulla sicurezza delle persone. Si tratta di sistemi che, pur utilizzando algoritmi di apprendimento, non incidono su decisioni cliniche o su ambiti regolati. In ambito sanitario possono includere, ad esempio, strumenti di gestione amministrativa, filtri antispam, analisi statistiche interne o algoritmi di raccomandazione di contenuti informativi. Per queste applicazioni l'AI Act non prevede obblighi specifici, ma incoraggia comunque l'adozione volontaria di buone pratiche di qualità, sicurezza e supervisione umana.



Figura 6 - Classificazione dei livelli di rischio previsti dal Regolamento (UE) 2024/1689 ("AI Act"). Il modello si basa su un approccio graduale, che distingue quattro categorie di rischio – inaccettabile, alto, limitato e minimo – con obblighi regolatori proporzionati all'impatto potenziale dei sistemi di intelligenza artificiale su diritti, sicurezza e libertà delle persone.

Per i sistemi ad **alto rischio**, l'AI Act impone una serie di **requisiti obbligatori** e una procedura di **valutazione di conformità** prima della loro immissione sul mercato. I requisiti chiave (Art. 8-15 del Regolamento) possono essere riassunti come segue:

- **Sistema di gestione del rischio:** il fornitore deve implementare un processo di gestione del rischio continuo, che copra analisi, valutazione e mitigazione dei rischi associati all'IA (includendo rischi di funzionamento errato, uso improprio, cybersecurity, ecc.).

- **Qualità dei dati:** i **dataset di addestramento, validazione e testing** devono essere di alta qualità in termini di rilevanza, rappresentatività, completezza e correttezza, per minimizzare distorsioni e risultati discriminatori. In particolare, va evitato che i bias nei dati portino a disparità di trattamento per gruppi protetti (*per età, genere, etnia, etc.*). Questo requisito riflette direttamente il principio OMS di equità e le tutele del diritto UE contro la discriminazione.
- **Tracciabilità e log:** i sistemi ad alto rischio devono essere progettati per **registrare automaticamente le loro operazioni (logging)** così da consentire tracciabilità e audit ex post dei risultati. Ad esempio, un CDS dovrà tenere traccia delle raccomandazioni fornite, parametri chiave, eventuali errori, in modo che in caso di incidente si possa capire cosa sia accaduto (analogia con la “scatola nera” degli aerei). Ciò è cruciale per accertare responsabilità e migliorare i sistemi.
- **Documentazione tecnica dettagliata:** il fornitore deve redigere un **fascicolo tecnico** con tutte le informazioni sul sistema, la sua progettazione, scopo, architettura, dati utilizzati, metriche di performance, ecc., necessarie alle autorità per valutarne la conformità. Questa documentazione è simile a quella richiesta per dispositivi medici, e deve essere mantenuta aggiornata. Nel caso di modifiche significative all’algoritmo, occorre aggiornare la documentazione e potenzialmente rifare la valutazione di conformità.
- **Informazioni e istruzioni d’uso:** chi fornisce un’IA ad alto rischio deve predisporre **informazioni chiare per l’utente finale** (il “*deployer*”). In campo sanitario, l’utente può essere il professionista o la struttura sanitaria che implementa il sistema. Le istruzioni devono spiegare lo scopo previsto, le prestazioni attese, i limiti noti del sistema, i requisiti di input, le condizioni ambientali, e anche dare indicazioni su come interpretare i risultati. Ad esempio, per un software di triage, il manuale deve dire che l’algoritmo stima un livello di urgenza, ma che questo è un supporto e *non sostituisce la valutazione clinica*, indicando magari come comportarsi in caso di conflitto tra algoritmo e giudizio umano. Inoltre, se pertinente, devono essere chiariti i **dati su cui è stato addestrato** (per es. “*addestrato su pazienti adulti, pertanto non validato in popolazioni pediatriche*”).

- **Supervisione umana:** l'AI Act richiede che i sistemi ad alto rischio siano progettati in modo da permettere ed agevolare una **adeguata supervisione umana** durante il loro utilizzo (Art. 14). In pratica, significa assicurare che l'IA non prenda decisioni in autonomia senza possibilità di intervento umano. Il grado e la forma di supervisione possono variare: ad esempio, può consistere nel mantenere un operatore **"in the loop"** (*che conferma ogni decisione prima dell'attuazione*) o **"on the loop"** (*pronto a intervenire e correggere*). Nel settore medicale, questo principio era già consolidato: un CDS deve lasciare al medico la decisione finale. Il Regolamento ora lo impone esplicitamente: design dell'interfaccia e training devono far sì che gli utenti umani comprendano le **capacità e i limiti del sistema** e possano giudicare quando deviare dalle sue raccomandazioni. Inoltre, in alcuni casi potrebbe essere previsto che l'IA abbia meccanismi di **spegnimento** o **fallback** se rileva malfunzionamenti, in modo che il controllo torni all'operatore.
- **Robustezza, accuratezza e sicurezza:** si esige un **elevato livello di robustezza e accuratezza** dei sistemi (Art. 15). Ciò implica che il sistema deve comportarsi in modo affidabile in diverse condizioni operative, gestire input errati o inattesi in sicurezza (*es. segnalando l'errore invece di fornire output fuorvianti*) e mantenere le prestazioni dichiarate entro range accettabili. La **cybersecurity** rientra qui: il sistema deve essere progettato per resistere ad attacchi e accessi non autorizzati, considerati anch'essi rischi per la sicurezza (*pensiamo a cosa significherebbe hackerare un algoritmo che supporta prescrizioni terapeutiche*). In ambito sanitario, la robustezza tocca anche la questione delle **versioni "statiche"** vs **"dinamiche"**: se un modello di IA continua ad apprendere in uso (*adaptive AI*), ciò potrebbe comportare variazioni non controllate di performance; i requisiti di robustezza e monitoraggio implicano qui limitazioni o controlli stringenti su tali approcci (*tanto che molti fornitori opteranno per modelli statici aggiornati periodicamente con nuove versioni sottoposte a valutazione di conformità*).

Oltre a questi requisiti tecnici, l'AI Act stabilisce un **sistema di governance e sorveglianza**: ogni sistema ad alto rischio dovrà passare per una **valutazione di conformità** prima dell'immissione sul mercato. Per molte applicazioni (*specialmente se rientranti anche come dispositivo medico*), questa valutazione coinvolgerà enti

terzi (*Notified Bodies* nel linguaggio dei dispositivi medici) che certificheranno la conformità ai requisiti essenziali dell'AI Act, rilasciando la marcatura CE specifica. Sarà istituito un **Registro UE dei sistemi di IA ad alto rischio**, dove ogni sistema dovrà essere registrato prima di entrare in servizio, incrementando la trasparenza. Le autorità di sorveglianza del mercato (*in Italia probabilmente il Ministero impresa/MiSE o organi designati*) avranno poteri di verificare la conformità, chiedere informazioni e irrogare sanzioni in caso di violazioni.

È importante sottolineare che, per evitare doppiopioni, l'AI Act contiene disposizioni di **allineamento con altre normative esistenti**. Nel caso dei dispositivi medici e IVD, che già seguono le normative MDR/IVDR, l'AI Act prevede *meccanismi di coordinamento*: ad esempio, la valutazione di conformità AI può essere svolta contestualmente a quella di dispositivo (*idealmente dallo stesso Notified Body*), e per alcuni requisiti ci si può basare su evidenze già fornite per la marcatura medica. Tuttavia, l'industria medtech ha espresso preoccupazioni su **potenziali sovrapposizioni** e **carichi aggiuntivi**: un produttore di software medico AI dovrà soddisfare sia MDR che AI Act, il che significa gestire due quadri regolatori complessi in parallelo (MedTech Europe, 2025). MedTech Europe (*associazione di settore*) ha chiesto un'implementazione coerente che eviti ritardi nell'accesso a tecnologie salvavita a causa di iter burocratici duplicati. In particolare, è stata proposta un'estensione dei tempi di applicazione per i sistemi già coperti da MDR/IVDR, per dare modo di elaborare standard armonizzati e linee guida specifiche senza interrompere l'innovazione. Il Regolamento, come da timeline ufficiale, concede in effetti un periodo più lungo per i sistemi di Allegato II (*embedded in prodotti regolamentati*) – fino al 2027 – proprio per allineare questi aspetti.

Dal punto di vista **dell'ambito sanitario**, l'AI Act impatterà fortemente: la maggior parte dei software di IA clinica sarà high-risk e dovrà quindi passare per queste forche caudine di qualità e compliance. Questo è positivo per i pazienti (*garanzie di efficacia e sicurezza*) ma pone sfide per produttori e per i sistemi sanitari che dovranno adeguarsi. Si prevedono ad esempio requisiti di **auditabilità dei CDS ospedalieri**: un ospedale che sviluppa internamente un algoritmo per, poniamo, ottimizzare le agende chirurgiche, dovrà esso stesso rispettare o attestare i requisiti

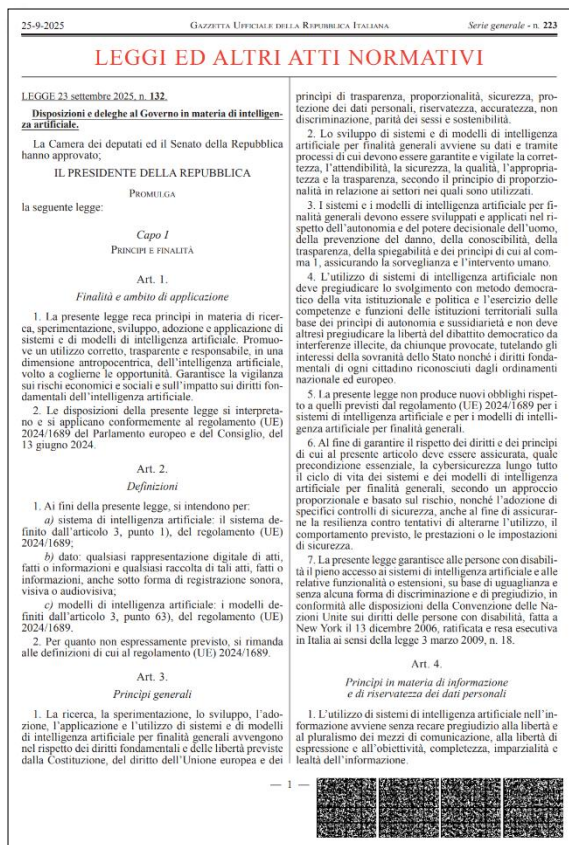
se l'algoritmo rientra nel campo del Regolamento (*anche un ospedale può essere "fornitore" ai sensi dell'AI Act se immette in servizio un sistema per uso interno professionale*). Ciò incoraggerà le strutture a collaborare con fornitori specializzati o consorzi che possano curare la compliance tecnica. Inoltre, dal lato **professionale**, il personale sanitario vedrà formalizzati alcuni obblighi: gli **"utilizzatori"** (*users deployer*) di sistemi ad alto rischio – come un ospedale o un medico che li adotta – avranno doveri quali utilizzare i sistemi secondo le istruzioni, monitorare le performance e segnalare eventuali incidenti gravi alle autorità (European Commission, 2024). Questo richiama un parallelo con la farmacovigilanza e la vigilanza dispositivi: anche per l'IA si creerà una cultura di segnalazione e apprendimento dagli errori.

Un aspetto particolare è l'attenzione a **trasparenza verso gli utenti finali** anche per sistemi non high-risk: l'AI Act introduce obblighi di trasparenza di informazioni (*disclosure*) per IA a *rischio limitato* come chatbot e deepfake. Ad esempio, se un paziente interagisce con un chatbot per informazioni sanitarie, ha diritto a sapere che sta parlando con un sistema automatizzato e non con un essere umano. Questo per preservare la fiducia e il *consenso informato* nell'uso di tali strumenti. Nei contesti clinici, è pensabile che linee guida etiche spingano a informare i pazienti quando le loro cure sono influenzate da un algoritmo (*e la legge italiana 132/2025 lo rende addirittura obbligo, cfr. sottocapitolo 2.2.3*).

Per concludere, l'AI Act fornisce un framework completo che **traduce in obblighi legali molti dei principi discussi dall'OMS**: gestione del rischio, qualità dei dati, trasparenza, supervisione umana, ecc. Per il personale sanitario, questo quadro significa operare in un ambiente dove le soluzioni di IA che incontreranno (*almeno quelle lecite*) saranno state certificate e dovranno mostrare certe garanzie. Tuttavia, resta fondamentale il **loro ruolo attivo**: il regolamento richiede supervisione umana ed è proprio il personale sul campo che fungerà da *"guardiano"* finale. Ad esempio, un algoritmo di supporto in sala parto potrebbe indicare un allarme per rischio di complicità: l'ostetrica dovrà saper interpretare quell'output e mettere in atto controlli, ma anche avere la prontezza di giudizio per capire se si tratta di un *falso allarme* dovuto a un limite noto del modello (*informazione che deve esserle stata fornita nel training e manuale*). La normativa le garantisce che l'algoritmo è stato

addestrato e validato secondo certi standard, ma non esime dalla **vigilanza professionale**. Inoltre, se qualcosa va storto (*ad es. l'IA sbaglia e causa un danno*), il quadro normativo chiarisce che pazienti e professionisti hanno diritto a spiegazioni e rimedi – e qui i professionisti sanitari dovranno farsi parte attiva nel segnalare incidenti e collaborare nelle analisi post-evento. Questo implica una certa **alfabetizzazione digitale**: il personale sanitario del futuro prossimo dovranno avere familiarità almeno basilare con concetti come *log di sistema, performance metriche, bias*, per poter partecipare efficacemente a questi processi di feedback e miglioramento continuo richiesti dalla legge.

2.2.3 La legge italiana 23 settembre 2025 n. 132: principi e implicazioni per l'uso dell'IA in sanità



L'Italia è stata tra i primi Paesi in Europa a dotarsi di una legislazione nazionale organica sull'Intelligenza Artificiale, anticipando in parte l'applicazione completa dell'AI Act europeo. La **Legge 23 settembre 2025, n. 132** - "*Disposizioni e deleghe al Governo in materia di tecnologie basate sull'intelligenza artificiale e responsabilità derivanti dal loro impiego*" - rappresenta il **primo intervento legislativo italiano dedicato all'IA**, entrato in vigore il 10 ottobre 2025 (Testa, 2025). Tale legge si muove nel solco dei principi

del Regolamento UE (*con cui dovrà armonizzarsi*) e della Costituzione italiana, e tocca vari settori (*pubblica amministrazione, giustizia, professioni, industria*). Di particolare rilievo per questa tesi sono gli **articoli 7-10**, che affrontano specificamente l'impiego dell'IA in ambito sanitario, delineando principi e cautele

per l'uso di sistemi di IA nella diagnosi, cura, prevenzione e gestione dei servizi sanitari. Analizziamo sinteticamente questi articoli:

Art. 7 - Principi fondamentali per l'utilizzo dell'IA in sanità. L'IA viene inquadrata dalla legge come **strumento di supporto** al servizio della persona e della medicina, con confini d'uso ben delimitati per evitare derive lesive dei diritti. In particolare, l'art.7 stabilisce che *“nei contesti sanitari l'intelligenza artificiale può essere utilizzata a supporto delle attività di diagnosi, prevenzione, cura e riabilitazione, ma non può in nessun caso sostituire il giudizio clinico del professionista sanitario”*. Questo principio per legge ribadisce la centralità insostituibile del medico (o dell'operatore sanitario qualificato) nel processo decisionale clinico, in linea con il principio dell'autonomia umana promosso dall'OMS e con le disposizioni dell'AI Act in materia di supervisione umana, configurando un modello di *Human-in-the-Loop* (o, più propriamente, di *Human-in-Command*), nel quale è esplicitamente **esclusa la delega decisionale totale al sistema di IA**. Inoltre, l'art.7 vieta espressamente che l'uso dell'IA possa determinare *“disparità di trattamento tra i pazienti”*. Non si possono usare algoritmi che, direttamente o indirettamente, *discriminino* pazienti in base a caratteristiche personali irrilevanti (*età, sesso, etnia, condizione socio-economica, stato di disabilità, etc.*). Si afferma così il principio di **equità algoritmica**, o *“algorithmic fairness”*, come principio giuridico: i sistemi di IA devono essere progettati e addestrati su dati di qualità evitando distorsioni che creino discriminazioni. Questo richiama gli obblighi di **quality/data governance** dell'AI Act, ma la legge italiana lo pone come principio di ordine pubblico: nella sanità italiana, l'IA non potrà essere uno strumento per *selezionare* chi merita cure su basi arbitrarie - la tutela dell'art.32 Cost. (*diritto alla salute*) e dell'art.3 (*eguaglianza*) viene così estesa all'ambito digitale.

L'art.7 introduce anche il **“diritto all'informazione”** del paziente circa l'impiego di IA nel proprio percorso di cura. In particolare, prevede che il paziente abbia il diritto di essere **consapevolmente informato** se una diagnosi o decisione terapeutica è stata coadiuvata da sistemi di IA, comprendendone le finalità, i limiti e il grado di affidabilità. Questo recepisce nell'ordinamento il concetto di *consenso informato digitale*: così come il paziente va informato di rischi/benefici di una procedura medica, va anche informato dell'uso di strumenti IA e delle loro caratteristiche (ad

es. “questo software aiuta a identificare anomalie, ha un’accuratezza del X% e non sostituisce il parere del medico”). Tale diritto all’informazione rafforza la posizione del paziente, permettendogli di *esercitare controllo* e partecipare attivamente al processo (*coerentemente con il GDPR art.22 sul diritto di non essere soggetti a decisioni completamente automatizzate senza spiegazioni*). Per gli operatori sanitari ciò implica un dovere di trasparenza e comunicazione: ad esempio, un chirurgo o un radiologo dovrà menzionare nel referto che è stato utilizzato un algoritmo di analisi delle immagini come supporto, e spiegare al paziente in termini comprensibili cosa questo comporta.

Un ultimo comma dell’art.7 incentiva l’uso dell’IA per finalità inclusive: promuove lo **sviluppo di soluzioni di IA a sostegno delle persone con disabilità**, richiedendo che tali sistemi siano progettati secondo criteri di accessibilità e inclusione. Questo punto sottolinea come la tecnologia possa essere strumento di empowerment - ad esempio sistemi di IA per facilitare la comunicazione di pazienti con disabilità verbali, protesi intelligenti, ausili robotici - in linea con la Convenzione ONU sui diritti delle persone con disabilità. La legge incoraggia quindi la **responsabilità sociale dell’innovazione**, il digitale dev’essere un fattore di emancipazione e non di esclusione per le categorie fragili. Ciò riflette una visione di **welfare digitale** in cui l’IA integra e potenzia i servizi assistenziali, con particolare riguardo anche alla telemedicina, teleriabilitazione, assistenti virtuali per anziani, ecc., tutti ambiti in cui infermieri e altre professioni sanitarie avranno ruoli di primo piano.

Art. 8 - Trasparenza dei processi decisionali con IA in sanità. Questo articolo rafforza il tema della **trasparenza**: dispone che ogni impiego di algoritmi o sistemi automatizzati da parte di strutture sanitarie, pubbliche o private, **deve essere accompagnato da informative chiare e verificabili** rivolte agli utenti (*pazienti*) e, su richiesta, alle autorità competenti. In altre parole, se un ospedale adotta un sistema di IA (*ad esempio, per lo smistamento di esami radiologici urgenti*), ciò va dichiarato e spiegato. La trasparenza non si esaurisce nella *comunicazione dell’uso*, ma include la possibilità per il paziente o per l’autorità di **accedere a informazioni sul funzionamento del modello, sui dati utilizzati e sulle modalità con cui è garantita la supervisione umana**. Questo aspetto ha un impatto notevole, in quanto implica che le strutture sanitarie devono poter fornire *spiegazioni tecniche* in caso di richiesta

(ad esempio, se un paziente contesta un risultato), l'ospedale deve poter dare dettagli (per quanto semplificati) del perché l'algoritmo ha prodotto quel risultato. È un onere che ricade sui **fornitori dei sistemi** (che dovranno dare alle strutture la documentazione necessaria) e sulle strutture stesse (che dovranno formare personale in grado di gestire tali richieste). Dal lato pratico, l'art.8 comporta che in ogni processo diagnostico-terapeutico in cui interviene l'IA "deve sempre essere previsto l'intervento di un medico responsabile, in grado di interpretare criticamente gli output del sistema e assumerne la responsabilità clinica". Questo passaggio è fondamentale: si vieta, ad esempio, che un software produca direttamente un referto consegnato al paziente senza visto medico, oppure che un piano di trattamento generato dall'IA venga applicato senza l'approvazione di un clinico. Viene **riafferma la centralità del professionista sanitario** e la sua **accountability**, proteggendolo al contempo dal rischio di "deresponsabilizzazione" di fronte ad algoritmi opachi. In altri termini, la legge tutela sia il paziente (che sa che c'è comunque un umano responsabile) sia il medico/infermiere/altro professionista sanitario (che non può essere messo nella posizione di dover seguire ciecamente un sistema che non capisce, perché ha diritto a informazioni su di esso). Questo articolo, quindi, istituzionalizza a livello nazionale quel concetto di **Human-in-the-loop** e **contestabilità** delle decisioni automatizzate che appare nei principi OMS e nell'AI Act. Per infermieri o altre professioni sanitarie, significa che, ad esempio, se utilizzano un sistema di allerta automatica in reparto (tipo *early warning system basato su IA*), dovranno comunque valutare personalmente il paziente e **validare** o meno l'allarme prima di agire; e qualora l'algoritmo facesse errori, non potranno giustificarsi dicendo "è colpa della macchina" perché rimane in capo a loro il dovere di vigilanza professionale, ma allo stesso tempo la struttura dovrà metterli a conoscenza di come la macchina funziona per permettere loro questa valutazione critica.

Art. 9 - Obblighi di sicurezza e responsabilità. Questo articolo impone che i sistemi di IA utilizzati in ambito sanitario rispettino **standard elevati di affidabilità, tracciabilità e protezione dei dati sensibili**. Si collega quindi ai requisiti tecnici: un sistema per essere impiegato in sanità italiana deve essere robusto e sicuro. Viene esplicitato che la norma è "coerente con il Regolamento europeo sull'IA", e prevede

che sia i fornitori (chi sviluppa e vende l'IA) sia gli utilizzatori (le strutture sanitarie, i professionisti) **adottino misure di cybersecurity adeguate** per prevenire manipolazioni, errori diagnostici o fughe di dati clinici. Le strutture sanitarie devono assicurarsi che il software di IA sia installato in ambiente sicuro, aggiornato, protetto da accessi indebiti, e che eventuali output siano verificati; i fornitori devono garantire protezioni contro manomissioni (*ad es. proteggere i modelli da attacchi adversarial che potrebbero generare diagnosi errate*). La legge delinea anche una **responsabilizzazione "differenziata" di progettisti, gestori e operatori sanitari**. Ciò significa riconoscere che la catena di responsabilità è condivisa: il produttore risponde di difetti dello strumento, l'ente che lo implementa risponde di una corretta integrazione e formazione del personale, l'operatore sanitario risponde dell'uso diligente. Si richiama la logica del "*safety by design*", ovvero integrare la sicurezza fin dalla progettazione e utilizzo del sistema, non limitarsi a controlli ex post. Un punto innovativo è l'accenno a un **regime di responsabilità per danni da IA ispirato a quello per prodotti difettosi**, in caso di danno al paziente causato da decisioni o azioni di IA, si prospetta che la prova dell'adeguatezza tecnica del sistema e del suo uso corretto sarà decisiva per stabilire la responsabilità. In altre parole, se un paziente subisce un danno (*es. diagnosi tardiva perché un algoritmo ha mancato un segno e il medico si è fidato*), si valuterà:

- il sistema era difettoso (*errore oltre soglia, non segnalato nei limiti*)?
- Il professionista lo ha usato secondo istruzioni?

Questa impostazione prepara il terreno a una disciplina specifica, anche considerando che l'UE sta lavorando ad aggiornare la *Product Liability Directive* per coprire l'IA. Per gli operatori sanitari, ciò implica la necessità di seguire protocolli e istruzioni d'uso alla lettera: se un infermiere ignora un warning del sistema senza ragionevole motivazione e succede un evento avverso, potrebbe esserne chiamato a rispondere; viceversa, se segue le linee guida e comunque c'è un errore dell'IA non prevedibile, la responsabilità potrà ricadere sul produttore del software. La legge mira dunque a chiarire e incentivare comportamenti diligenti lungo tutta la filiera, evitando la terra di nessuno delle responsabilità.

Art. 10 - Diritti degli utenti e dei pazienti riguardo all'IA. Questo articolo chiude il quadro “*garantista*” delineando una serie di **diritti individuali** nei confronti dei sistemi di IA in sanità. Viene sancito che ogni persona ha **diritto di sapere** e comprendere come l'IA ha inciso su una decisione sanitaria che la riguarda, di **chiedere chiarimenti** o **una revisione umana della decisione automatizzata** e, se necessario, di **contestare l'esito o ottenere un risarcimento** in caso di errore. In sostanza, se un paziente ritiene che un algoritmo abbia influenzato negativamente la sua cura (*ad es. è stato classificato a minor priorità per un intervento e crede sia per via di un bias*), ha il diritto di chiedere spiegazioni all'azienda sanitaria, ottenere che un medico rivaluti la sua situazione senza l'interferenza dell'algoritmo, e se ha subito un danno provabile, chiederne conto legalmente. Questa norma **estende al contesto digitale le garanzie già previste dal GDPR** (*artt. 13-15,22: diritto alla trasparenza, accesso ai dati, opposizione a decisioni automatizzate*) e le cala specificamente nella tutela del diritto alla salute. La ratio è impedire che l'innovazione tecnologica si traduca in *opacità decisionale* o in *de-responsabilizzazione* di fronte al cittadino. In ambito sanitario c'è sempre stato il principio per cui il paziente può chiedere una *second opinion* o contestare una decisione: la legge chiarisce che questo vale anche se la decisione coinvolge l'IA, anzi, a maggior ragione, il paziente dev'essere protetto da eventuali “*scaricabarile*” su un computer. In termini di sistema, questo articolo spinge le strutture a predisporre canali e procedure per gestire tali richieste (*ad es. un comitato interno per riesaminare casi controversi dove era coinvolta IA, o policy per informare i pazienti in cartella su quali decisioni sono state automatizzate*). Per il personale sanitario, ciò significa che dovranno essere pronti a interfacciarsi con pazienti che chiedono spiegazioni. Insomma, il personale sanitario diventa anche *garante* di questi diritti sul campo, essendo il front-end del sistema verso i cittadini.

Nel complesso, gli **articoli 7-10 della Legge 132/2025** dipingono una visione dell'IA come “**alleata**” della medicina e **non sua sostituta**, incardinata in una etica pubblica della cura basata sulla centralità della persona. La legge traduce in norme vincolanti molti dei principi visti: l'autonomia decisionale del clinico, la non discriminazione, la trasparenza spiegabile, la sicurezza by design, l'accountability e il diritto alla contestazione. Per realizzare questa visione, sarà cruciale l'emanazione di **decreti**

attuativi che definiscano standard tecnici chiari, procedure di controllo efficaci e strumenti di formazione per il personale sanitario. La stessa legge delega il Governo ad adottare decreti in vari ambiti: dalla definizione dei poteri di vigilanza delle autorità competenti (*in sinergia con l'AI Act*), alla creazione di **sandbox** normative per testare IA in sanità (*Art. 3 e 5 della legge introducono la possibilità di aree di sperimentazione protette, gestite dalle Autorità per l'IA, con regole specifiche per il settore sanitario per favorire innovazione sicura*). Inoltre, la legge tocca aspetti di **governance nazionale dell'IA**, individua nella Presidenza del Consiglio il soggetto guida per la strategia IA, crea un Comitato di coordinamento, e designa due autorità nazionali principali - l'Agenzia per la Cybersicurezza Nazionale e l'Agenzia per l'Italia Digitale - competenti anche per l'IA in sanità. Essa prevede cooperazione con le autorità esistenti come il Garante Privacy e chiarisce che alcune autorità settoriali (*ad es. AIFA per il farmaco, Ministero Salute per dispositivi*) manterranno le loro competenze, ma in un quadro coordinato. Ciò significa che in Italia si punta a un approccio *interdisciplinare*: la sicurezza cibernetica, la digitalizzazione, la protezione dati e la regolazione sanitaria dovranno muoversi all'unisono per governare l'IA.

Un altro elemento importante è la norma che **facilita la ricerca scientifica con dati sanitari**, permettendo di usare la base giuridica dell'**interesse pubblico** per trattare dati a fini di ricerca in ambito IA sanitario. Questo rimuove alcune incertezze del GDPR e può accelerare progetti di ricerca su ML applicato alla medicina (*garantendo però il rispetto di principi costituzionali e misure a tutela dei dati*). Ciò è visto come un vantaggio competitivo per il sistema Paese in termini di innovazione sanitaria.

2.3 Implicazioni operative per i professionisti sanitari

La Legge 132/2025, insieme al quadro regolatorio dell'Unione Europea e ai principi OMS, incide direttamente su come i professionisti sanitari, interagiranno con l'IA nei loro ambienti di lavoro. Riassumiamo le principali implicazioni:

- **Ruolo di supporto, non sostitutivo:** La normativa ribadisce che l'IA è uno **strumento di supporto**. Pertanto, il personale potrà avvalersi di sistemi IA, ma **non verranno rimpiazzati** da essi. Anzi, permane in capo a loro la responsabilità di supervisionare e validare le indicazioni della macchina. Questo dovrebbe

rassicurare i professionisti che la “*decisione finale*” resta la loro. Ciò tutela anche la professionalità: IA come estensione delle competenze, non come sostituzione.

- **Formazione e competenze digitali:** Affinché possano davvero supervisionare l’IA in modo appropriato, i professionisti sanitari dovranno acquisire nuove competenze in ambito digitale. Le normative non lo esplicitano, ma è un corollario: l’OMS sottolinea la necessità di “*digital literacy or retraining*” per milioni di operatori sanitari che dovranno interagire con macchine che possono sfidarne il processo decisionale (World Health Organization, 2021). In pratica, i curricula formativi dovranno includere elementi di informatica medica, interpretazione di output di IA, gestione dei dati, principi etici e normativi sull’IA. È probabile che linee guida ministeriali e ordinistiche raccomanderanno corsi di aggiornamento specifici, dato che usare un dispositivo di IA senza comprenderne bene limiti e logica espone a rischi clinici e legali. Questa esigenza è riconosciuta implicitamente anche dall’AI Act, che con la **supervisione umana** richiede che gli utenti siano “*adeguatamente informati delle capacità e limiti del sistema*”: ciò avverrà tramite manuali, ma anche tramite training utente forniti dalle aziende.
- **Coinvolgimento nella co-progettazione e validazione:** La letteratura e le policy incoraggiano il coinvolgimento degli operatori sanitari **nella fase di sviluppo e implementazione** delle soluzioni di IA (approccio *human-centered design*). La legge italiana indirettamente spinge in questa direzione introducendo *sandbox sanitarie* dove innovatori, clinici e regolatori collaborano per testare soluzioni. I professionisti sanitari, essendo a diretto contatto con i pazienti e i flussi operativi, possono fornire un contributo fondamentale: possono segnalare requisiti funzionali, potenziali criticità pratiche e suggerire migliorie per adattare l’IA al contesto reale. Il *coinvolgimento proattivo* dei professionisti nei progetti IA è anche una garanzia di maggiore accettazione e fiducia verso gli strumenti stessi.
- **Osservanza di protocolli e istruzioni d’uso:** Con strumenti IA certificati, verranno fornite linee guida operative specifiche. Il personale sanitario dovrà attenersi scrupolosamente a tali **protocolli**, perché la conformità all’uso previsto è sia un obbligo deontologico sia una protezione da responsabilità. La legge richiama che *l’uso corretto è decisivo in caso di contenzioso*, pertanto seguire la corretta

procedura rappresenta la miglior difesa. Le direzioni sanitarie dovranno quindi aggiornare le **procedure interne** includendo le fasi di intervento umano e di gestione degli output IA, e il personale dovrà documentare in cartella il proprio intervento (es: *“alert sistema EarlyWarning = 7, paziente rivalutato, parametri confermati, medico avvisato”* oppure *“alert rivelatosi falso positivo, monitoraggio continuo mantenuto”*).

- **Tutela della privacy e sicurezza dei dati:** I professionisti sanitari saranno sempre più coinvolti nella raccolta e gestione di dati che alimentano sistemi di IA (*dati clinici inseriti in sistemi, dati dei pazienti caricati per analisi, ecc.*). Devono quindi mantenere alta l’attenzione a **privacy e sicurezza**: ad esempio, assicurarsi di usare solo piattaforme autorizzate dall’ospedale, non esportare dati sensibili in applicazioni non approvate, rispettare i principi di minimizzazione (*inserire nei sistemi solo i dati realmente necessari*). La normativa (GDPR, legge 132/2025 art.9) li tutela pretendendo dai fornitori standard elevati di sicurezza, ma l’anello finale resta l’utente: un errore come lasciare incustodito un terminale loggato con dati di pazienti, o copiare dati su dispositivi non sicuri, può vanificare protezioni tecniche. Pertanto, gli enti dovranno formare il personale sulle policy di cybersecurity relative ai nuovi strumenti (ad es., autenticazione a due fattori, cambio password, segnalazione immediata di incidenti informatici).
- **Comunicazione e relazione col paziente:** Infine, le regole di trasparenza e diritto all’informazione significano che **i professionisti sanitari dovranno spesso farsi carico di spiegare al paziente** in che modo un’IA è stata utilizzata nella sua cura e rassicurarlo sul fatto che ciò non compromette l’umanità e la qualità della cura e dell’assistenza. Le professioni assistenziali (*ad es. infermieri ed ostetriche*), tradizionalmente più vicine al paziente nel quotidiano, dovranno saper tradurre concetti tecnici in parole semplici: ad esempio, spiegare che *“questo software ci aiuta a tenere d’occhio i suoi parametri e ci avvisa subito se qualcosa non va, ma stia tranquillo che noi controlliamo costantemente e qualsiasi allarme lo verifichiamo di persona”*. Oppure rispondere a dubbi: *“Perché mi hanno posticipato l’intervento?”* - *“Abbiamo un sistema che calcola delle priorità anche in base alla gravità; nel suo caso è stato valutato che un breve rinvio non*

comporta rischi, mentre abbiamo dovuto anticipare un caso più urgente. Se vuole maggiori dettagli, possiamo approfondire col chirurgo". Questo tipo di comunicazione empatica, supportata dalla trasparenza di cui la legge fa un diritto, sarà fondamentale per **mantenere la fiducia** dei pazienti nonostante l'ingresso delle macchine. Il personale infermieristico e ostetrico, in quanto fautore della caring profession (*professioni di cura*), avrà il compito di garantire che l'IA sia percepita come valore aggiunto e non come minaccia al rapporto umano di cura.

In conclusione, il **contesto regolatorio odierno** - dai principi dell'OMS, passando per il Regolamento UE, fino alla dettagliata legge italiana - costituisce un quadro di riferimento entro cui sviluppare e adottare soluzioni di IA in sanità in modo etico, sicuro e paziente-centrico. Le norme spingono verso un'IA "**affidabile per design**", in cui *efficacia tecnica e rispetto dei diritti* procedono di pari passo. In questo senso, la valutazione dell'**efficacia delle soluzioni di intelligenza artificiale in ambito ambulatoriale e chirurgico** non può limitarsi ai soli indicatori di performance o agli esiti clinici. Una tecnologia può essere considerata realmente efficace solo quando, oltre a dimostrare un miglioramento misurabile dei processi (*riduzione dei tempi di attesa, ottimizzazione delle risorse, incremento dell'accuratezza decisionale*), risulta accettabile e utilizzabile dai professionisti coinvolti, non introduce distorsioni algoritmiche, è sostenibile dal punto di vista organizzativo ed economico ed è pienamente conforme al quadro regolatorio vigente.

I capitoli successivi analizzeranno in modo sistematico la letteratura disponibile sulle applicazioni di IA nei setting ambulatoriali e chirurgici, adottando questa prospettiva integrata.

CAPITOLO 3 - Metodi della Revisione

3.1 Disegno dello studio e formulazione del PICO

La presente revisione narrativa della letteratura è stata condotta seguendo un approccio strutturato, volto a identificare, selezionare e analizzare gli studi riguardanti l'impiego dell'intelligenza artificiale in ambito ambulatoriale e chirurgico, con particolare attenzione sia all'ottimizzazione delle risorse sanitarie sia alla valutazione degli esiti clinici e organizzativi. A supporto della ricerca è stato adottato il modello PICO (Popolazione, Intervento, Confronto, Outcome), che ha permesso di definire in maniera chiara il quesito di partenza e gli ambiti di interesse. Il quesito di ricerca è stato strutturato come segue:

- **P (Popolazione):** Studi che analizzano l'applicazione di soluzioni di intelligenza artificiale in ambito ambulatoriale e chirurgico.
- **I (Intervento):** Implementazione di tecnologie di intelligenza artificiale (machine learning, deep learning, NLP, computer vision, modelli predittivi, sistemi di supporto decisionale).
- **C (Confronto):** Metodi tradizionali di gestione e pianificazione delle risorse sanitarie, o assenza di soluzioni di IA.
- **O (Outcome):**
 1. **Ottimizzazione delle risorse** → efficienza organizzativa, riduzione tempi/costi, miglior allocazione posti letto/sale operatorie/agenda ambulatoriale.
 2. **Valutazione degli esiti** → outcome clinici e organizzativi: accuratezza diagnostica, riduzione complicanze, riduzione tempi di attesa, soddisfazione del paziente, sicurezza, qualità percepita.

3.2 Ricerca su basi di dati tradizionali (PubMed)

La ricerca bibliografica è stata effettuata principalmente su PubMed, applicando i seguenti criteri di selezione:

- **Periodo di pubblicazione:** ultimi 10 anni (2015–2025).

- **Lingua:** inglese, spagnolo e italiano.
- **Tipo di studi:** studi primari, revisioni sistematiche o narrative, report di enti sanitari internazionali e paper scientifici peer-reviewed.
- **Contesto:** sanità umana (sono stati esclusi gli studi relativi alla sanità veterinaria).

Le stringhe di ricerca elaborate, distinte per ambito e outcome, sono riportate integralmente:

A. Ambulatoriale (outpatient, risorse, scheduling, outcome)

("Artificial Intelligence"[Mesh] OR "Machine Learning"[Mesh] OR "Deep Learning"[tiab] OR "Neural Networks, Computer"[Mesh] OR "Natural Language Processing"[tiab] OR "Clinical Decision Support Systems"[Mesh]) AND ("Ambulatory Care"[Mesh] OR "Outpatients"[Mesh] OR "Outpatient Clinics, Hospital"[Mesh] OR "Primary Health Care"[Mesh] OR outpatient[tiab] OR "ambulatory care"[tiab]) AND ("Appointments and Schedules"[Mesh] OR "Resource Allocation"[Mesh] OR "Efficiency, Organizational"[Mesh] OR "Waiting Lists"[Mesh] OR "Patient Acceptance of Health Care"[Mesh] OR schedule*[tiab] OR workflow[tiab] OR optimization[tiab] OR "no-show"[tiab] OR "nonattendance"[tiab] OR throughput[tiab] OR "capacity planning"[tiab]) AND ("Treatment Outcome"[Mesh] OR "Patient Outcome Assessment"[Mesh] OR outcome*[tiab] OR effectiveness[tiab] OR "clinical outcome*" [tiab] OR "patient safety"[Mesh] OR "quality of health care"[Mesh] OR "quality improvement"[tiab]) NOT (animals[Mesh] NOT humans[Mesh]) AND ("2015/01/01"[Date - Publication] : "3000"[Date - Publication]) AND (English[lang] OR Italian[lang] OR Spanish[lang])*

B. Chirurgico / Sala operatoria (OR management, outcome)

("Artificial Intelligence"[Mesh] OR "Machine Learning"[Mesh] OR "Deep Learning"[tiab] OR "Neural Networks, Computer"[Mesh] OR "Natural Language Processing"[tiab] OR "Clinical Decision Support Systems"[Mesh]) AND ("Operating Rooms"[Mesh] OR "Surgical Procedures, Operative"[Mesh] OR "Perioperative Care"[Mesh] OR "Operating Room Nursing"[Mesh] OR "Operating Room Information Systems"[tiab] OR "operating room"[tiab] OR perioperat[tiab]) AND ("Resource Allocation"[Mesh] OR "Efficiency, Organizational"[Mesh] OR "Appointments and*

Schedules[Mesh] OR *schedule**[tiab] OR "case duration"[tiab] OR "turnover time"[tiab] OR *cancellation**[tiab] OR *utilization*[tiab] OR *throughput*[tiab] OR *block**[tiab]) AND ("Treatment Outcome"[Mesh] OR "Patient Outcome Assessment"[Mesh] OR *outcome**[tiab] OR *effectiveness*[tiab] OR "clinical outcome*" [tiab] OR "patient safety"[Mesh] OR "quality of health care"[Mesh] OR "quality improvement"[tiab]) NOT (*animals*[Mesh] NOT *humans*[Mesh]) AND ("2015/01/01"[Date - Publication] : "3000"[Date - Publication]) AND (*English*[lang] OR *Italian*[lang] OR *Spanish*[lang])

C. Focus “solo scheduling/ottimizzazione” (ambulatoriale o chirurgico, con outcome)

("Artificial Intelligence"[Mesh] OR "Machine Learning"[Mesh] OR "Deep Learning"[tiab] OR "Reinforcement Learning"[tiab] OR "Predictive Analytics"[tiab]) AND ("Appointments and Schedules"[Mesh] OR *scheduling*[tiab] OR *reschedul**[tiab] OR *optimization*[tiab] OR "resource allocation"[tiab] OR "capacity planning"[tiab]) AND (*outpatient**[tiab] OR "ambulatory care"[tiab] OR "operating room"[tiab] OR *perioperat**[tiab] OR *surgical*[tiab]) AND ("Treatment Outcome"[Mesh] OR "Patient Outcome Assessment"[Mesh] OR *outcome**[tiab] OR *effectiveness*[tiab] OR "clinical outcome*" [tiab] OR "patient safety"[Mesh] OR "quality of health care"[Mesh] OR "quality improvement"[tiab]) NOT (*animals*[Mesh] NOT *humans*[Mesh]) AND ("2015/01/01"[Date - Publication] : "3000"[Date - Publication]) AND (*English*[lang] OR *Italian*[lang] OR *Spanish*[lang])

D. Combinata “AI + processi organizzativi” (ampia, screening iniziale)

("Artificial Intelligence"[Mesh] OR "Machine Learning"[Mesh] OR "Deep Learning"[tiab] OR "Neural Networks, Computer"[Mesh] OR "Natural Language Processing"[tiab]) AND ("Efficiency, Organizational"[Mesh] OR "Resource Allocation"[Mesh] OR "Health Services Administration"[Mesh] OR *workflow*[tiab] OR "process improvement"[tiab] OR *optimization*[tiab]) AND (*outpatient**[tiab] OR "ambulatory care"[tiab] OR "operating room"[tiab] OR *perioperat**[tiab] OR *surgical*[tiab]) AND ("Treatment Outcome"[Mesh] OR "Patient Outcome Assessment"[Mesh] OR *outcome**[tiab] OR *effectiveness*[tiab] OR "clinical outcome*" [tiab] OR "patient safety"[Mesh] OR "quality of health care"[Mesh] OR

"quality improvement"[tiab]) NOT (animals[Mesh] NOT humans[Mesh]) AND ("2015/01/01"[Date - Publication] : "3000"[Date - Publication]) AND (English[lang] OR Italian[lang] OR Spanish[lang])

3.3 Piattaforme di Intelligenza Artificiale per la ricerca (Consensus, Elicit, ChatGPT)

Per integrare la ricerca tradizionale, sono state utilizzate anche piattaforme di intelligenza artificiale specializzate nella ricerca scientifica: **Consensus** (funzione deep research-PRO), **Elicit** (funzione deep research) e **ChatGPT 5.0 Plus** (deep research). Per garantire uniformità e comparabilità dei risultati è stato utilizzato un unico prompt di ricerca, in lingua inglese, di cui si riporta anche la traduzione italiana.

“Conduct an in-depth literature search on the use of Artificial Intelligence — including machine learning, deep learning, neural networks, and other related technologies — for the optimization and planning of healthcare resources in outpatient and surgical settings. Focus on real-world applications, developed systems, and their impact in terms of efficiency, clinical outcomes, and resource utilization. The search should cover studies published in the last 10 years, in English, Spanish, or Italian, and include all types of articles (reviews, primary studies, research papers). Exclude studies related to veterinary healthcare”.

Traduzione in lingua italiana del Prompt:

“Effettua una ricerca bibliografica approfondita sull’impiego dell’Intelligenza Artificiale — includendo machine learning, deep learning, reti neurali e altre tecnologie — per l’ottimizzazione e la pianificazione delle risorse sanitarie in ambito ambulatoriale e chirurgico. Concentrati su applicazioni reali, sistemi sviluppati e sul loro impatto in termini di efficienza, esiti clinici e utilizzo delle risorse. La ricerca deve riguardare studi pubblicati negli ultimi 10 anni, in lingua inglese, spagnola o italiana, e includere tutte le tipologie di articoli (revisioni, studi primari, paper). Escludere gli studi riferiti alla sanità veterinaria.”

3.4 Gestione delle fonti e screening

La ricerca complessiva, condotta tramite le modalità sopra descritte, ha restituito **74 pubblicazioni scientifiche**. Per organizzare e analizzare il materiale è stata costruita una **griglia di estrazione dati degli studi** (*Allegato A*), nella quale a ciascuna pubblicazione è stato assegnato un **ID univoco** (*ST01, ST02 etc.*).

La griglia comprendeva i seguenti item:

- Titolo (in lingua originale e traduzione italiana).
- Autore/i
- Anno
- Paese
- Tipo di studio (RCT, osservazionale, revisione, etc.)
- Contesto (Ambulatoriale / Chirurgico)
- Specialità / setting
- Tecnologia IA (ML, DL, RL, NLP, Ibridi, altro)
- Algoritmo / Modello specifico
- Compito (Previsione, Ottimizzazione, Supporto decisionale, altro)
- Dataset (dimensione, fonte, real-world vs simulato)
- Outcome principali
- Metriche di valutazione (AUC, MAE, RMSE, % miglioramento, ecc.)
- Risultati chiave
- Benefici dichiarati
- Limiti dichiarati
- Riferimento bibliografico (DOI o PMID)

Successivamente è stata effettuata la lettura degli abstract, al fine di valutare l'inclusione o l'esclusione degli studi nella revisione narrativa. Questo processo ha consentito di ottenere una panoramica iniziale delle evidenze disponibili e di identificare i contributi più rilevanti in relazione agli obiettivi di ricerca. Dalla lettura dei 74 abstract inizialmente individuati, sono state incluse nella revisione 52 pubblicazioni. La composizione degli studi è risultata eterogenea, sia in termini di disegno metodologico sia di contenuti affrontati, riflettendo la multidimensionalità

del tema “*intelligenza artificiale e gestione delle risorse sanitarie*”. Dal punto di vista metodologico, la quota più rilevante è rappresentata dagli **studi osservazionali**, per lo più retrospettivi, spesso condotti su database ospedalieri di grandi dimensioni, con lo scopo di sviluppare e validare modelli predittivi basati su tecniche di *machine learning* (ST02, ST05, ST13, ST14, ST29, ST45, ST56, ST67). Sono presenti anche alcuni studi prospettici e before–after (ST11, ST32, ST40, ST65), che aggiungono elementi di maggiore solidità nella valutazione degli effetti reali. Accanto a questi, figurano diversi **studi metodologici e applicativi** orientati allo sviluppo di soluzioni algoritmiche, talvolta validate in contesti real-world o di simulazione (ST34, ST41, ST51, ST59, ST61, ST74).

Un secondo blocco significativo è costituito dalle **revisioni della letteratura**, sia sistematiche (alcune condotte secondo linee guida PRISMA, ST20, ST25, ST27, ST42, ST54) sia narrative (ST12, ST28, ST31, ST35, ST36, ST50, ST52, ST53, ST58). Queste pubblicazioni hanno permesso di sintetizzare lo stato dell’arte e di evidenziare prospettive applicative, lacune metodologiche e direzioni future di ricerca.

Sono inoltre presenti alcuni **studi qualitativi** (ST07) e contributi in forma di **editoriali, commentary o analisi critiche** (ST19, ST43, ST49, ST60, ST62, ST64), che pur non offrendo dati sperimentali, hanno fornito riflessioni contestuali e raccomandazioni utili all’implementazione dell’IA nei sistemi sanitari. Infine, è stato identificato un **protocollo di studio** (ST71), che pur non riportando risultati, rappresenta un segnale dell’interesse crescente verso studi comparativi rigorosi tra modelli predittivi.

Dal punto di vista delle **tematiche affrontate**, la maggior parte degli studi si concentra sullo **sviluppo di modelli predittivi** per migliorare la pianificazione e la gestione delle risorse (*no-show prediction, scheduling chirurgico, allocazione posti letto, turnover in sala operatoria*). Un numero consistente di pubblicazioni approfondisce invece gli aspetti di **ottimizzazione organizzativa** attraverso algoritmi matematici e sistemi decisionali AI-based, mentre altri contributi assumono una prospettiva più ampia, proponendo **analisi critiche, agende di ricerca o prospettive applicative**.

La distribuzione degli studi inclusi (n = 52) nei diversi setting è riportata nella tabella seguente:

Setting	N. studi	Codici Studio (ST)	Nota sintetica
Chirurgico	33	ST02, ST03, ST05, ST08, ST12, ST17, ST18, ST19, ST20, ST21, ST25, ST27, ST29, ST31, ST33, ST35, ST36, ST42, ST43, ST44, ST45, ST49, ST50, ST51, ST52, ST53, ST54, ST58, ST59, ST60, ST64, ST65, ST71	Prevalenza di studi osservazionali e revisioni; focus su scheduling OR, gestione sale operatorie, validazione/implementazione di modelli predittivi.
Ambulatoriale	15	ST01, ST06, ST07, ST11, ST13, ST14, ST15, ST32, ST34, ST41, ST46, ST56, ST61, ST67, ST74	Modelli ML per no-show, scheduling e gestione agende; alcuni studi prospettici/metodologici.
Generalista	3	ST10, ST28, ST62	Revisioni/commentari trasversali sull'IA in sanità senza focus di setting.
Misto	1	ST40	Studio prospettico applicato a entrambi i contesti.

La distribuzione degli studi inclusi (n = 52) in base alla tipologia metodologica è invece illustrata nel grafico seguente:

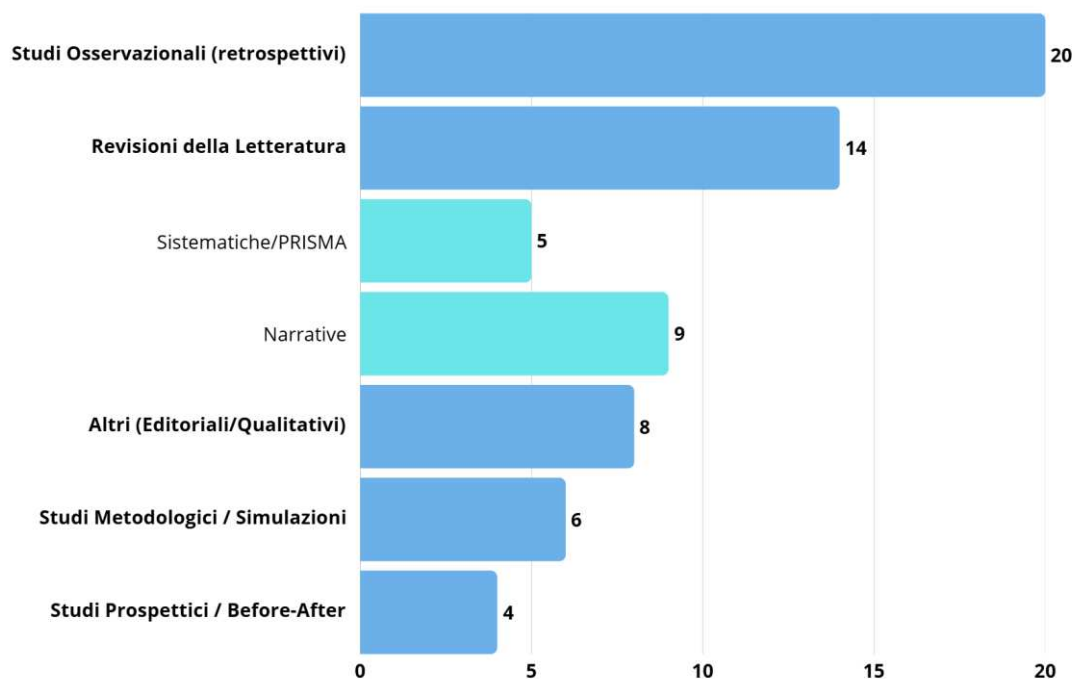


Figura 7 - Distribuzione metodologica dei 52 studi inclusi nella revisione.

I 52 studi inclusi delineano un quadro variegato e in evoluzione: da un lato l'ampia sperimentazione di modelli predittivi e metodologici, dall'altro la necessità di consolidare le evidenze attraverso studi prospettici, valutazioni di impatto sugli outcome clinici e implementazioni real-world.

Per quanto riguarda il **setting degli studi**, la distribuzione evidenzia una prevalenza di pubblicazioni in ambito **chirurgico** (33 studi), seguite da quelle in ambito **ambulatoriale** (15 studi). Sono inoltre presenti **3 contributi generalisti**, che affrontano l'IA in sanità senza un focus esclusivo, e **1 studio a cavallo tra i due contesti** (ST40), che ha considerato contemporaneamente setting ambulatoriali e chirurgici. Questa distribuzione mette in luce come l'IA stia trovando progressiva applicazione tanto nella chirurgia — dove la pressione organizzativa è particolarmente elevata — quanto negli ambulatori, contesti che pongono crescenti sfide di accesso, gestione dei flussi e riduzione delle liste d'attesa.

Parallelamente, **22 studi** sono stati esclusi dalla revisione. La maggior parte apparteneva all'ambito **generalista** (12 studi), trattando l'IA in sanità senza un collegamento diretto con i setting ambulatoriali o chirurgici, oppure con approcci troppo concettuali o esplorativi. Cinque studi riguardavano il contesto **chirurgico**, ma si trattava prevalentemente di editoriali, commenti o analisi prive di dati sperimentali. Ulteriori esclusioni hanno riguardato setting non pertinenti rispetto al quesito di ricerca: *extra-ospedaliero (1)*, *bed management (1)*, *cure primarie (1)*, *infermieristico (1)* e *gestione delle risorse umane (1)*.

Le motivazioni principali di esclusione sono state:

- **Assenza di dati originali** (commenti, editoriali, lettere, analisi concettuali).
- **Disegno non compatibile con gli obiettivi della revisione** (scoping review, horizon scan o articoli esplorativi).
- **Setting non pertinente** (extra-ospedaliero, infermieristico, gestione risorse umane).
- **Mancanza di applicazione reale di AI** o assenza di outcome clinici/organizzativi rilevanti.

La distribuzione degli studi esclusi è sintetizzata nella tabella seguente:

Setting	N. studi	Codici Studio (ST)	Nota sintetica
Chirurgico	5	ST04, ST16, ST23, ST24, ST47	Editoriali/commenti, mancanza di dati originali
Generalista (AI in sanità)	12	ST22, ST26, ST30, ST37, ST38, ST39, ST48, ST57, ST63, ST66, ST69	Analisi concettuali, scoping review, focus non mirato su risorse/esiti
Extra-ospedaliero	1	ST09	Setting non pertinente
Bed management	1	ST55	Setting non pertinente
Cure primarie	1	ST70	Setting non pertinente
Infermieristico	1	ST72	Setting non pertinente
Gestione risorse umane (generalista)	1	ST73	Focus diverso dagli obiettivi della revisione

Il percorso metodologico sopra descritto ha permesso di raccogliere, organizzare e valutare in modo strutturato le principali evidenze scientifiche disponibili sull'uso dell'intelligenza artificiale per l'ottimizzazione delle risorse e la valutazione degli esiti clinici e organizzativi in ambito ambulatoriale e chirurgico.

Nel capitolo successivo sono riportati i risultati derivanti dall'analisi delle pubblicazioni incluse, offrendo una sintesi delle evidenze disponibili.

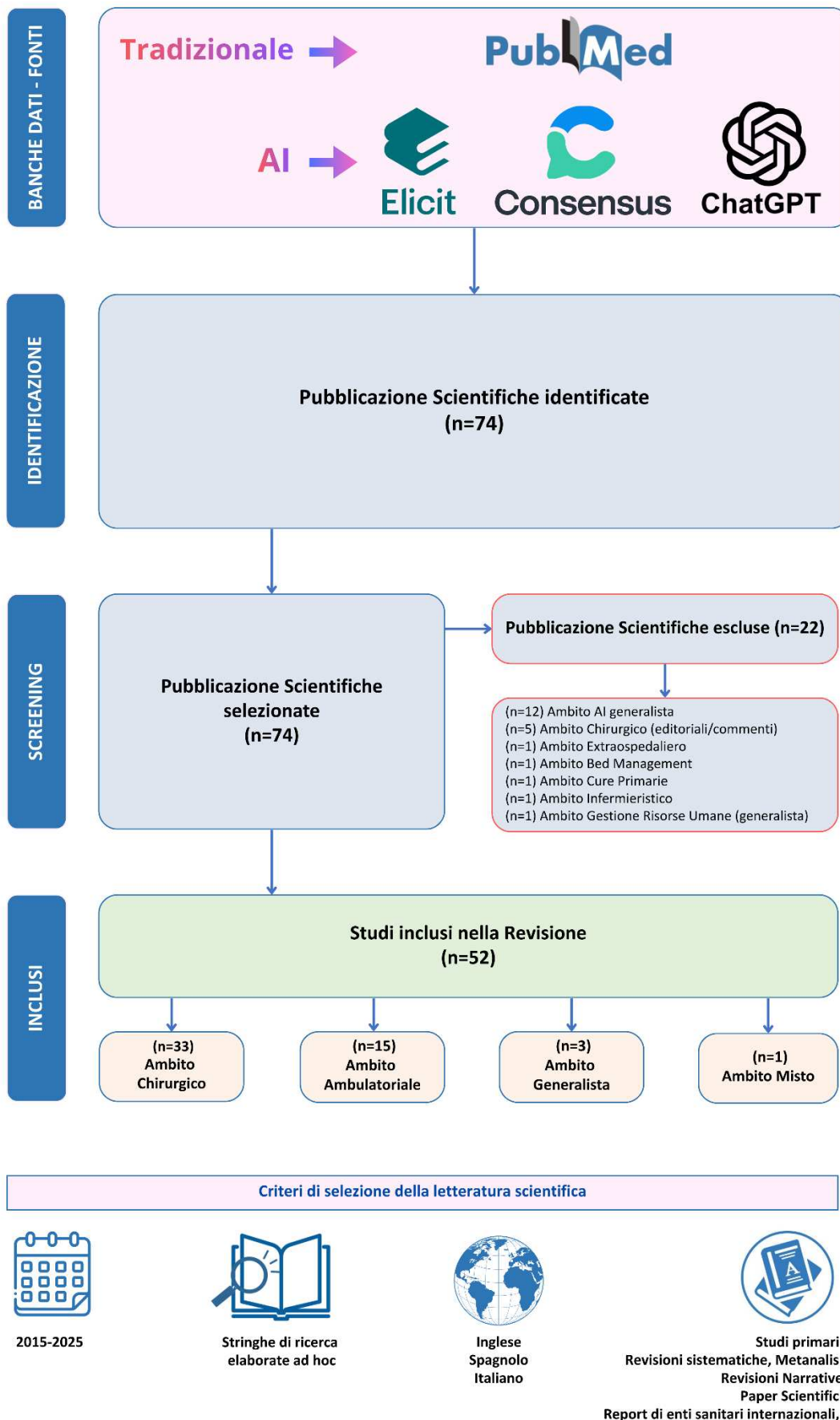


Figura 8 - Flow chart del processo di identificazione, screening e inclusione degli studi, con rappresentazione dei criteri utilizzati per la selezione della letteratura scientifica

CAPITOLO 4 - Risultati e Stratificazioni principali

4.1 Per ambiente di cura

4.1.1 Ambulatoriale/Outpatient (no show, triage/overbooking, capacity planning)

L'analisi dei 52 studi evidenzia un ampio ventaglio di applicazioni dell'IA in contesti ambulatoriali (*outpatient*), che spaziano dalla medicina generale alle specialità diagnostiche e cliniche. In questo ambiente, la sfida principale è gestire in modo efficiente le agende degli appuntamenti e mitigare le inefficienze dovute ai *no-show* (mancate presentazioni). Le strutture ambulatoriali operano spesso con capacità giornaliera limitata e alta variabilità di domanda. Un tasso significativo di pazienti (*anche intorno al 20%*) potrebbe non presentarsi agli appuntamenti, causando sprechi di tempo e risorse e allungando le liste d'attesa. Tali assenze comportano una perdita di efficienza stimata tra il 3% e il 14% dei ricavi e compromettono la continuità dell'assistenza. Di conseguenza, negli studi esaminati si riscontra un'attenzione particolare al miglioramento dell'aderenza agli appuntamenti e alla prevenzione di queste inefficienze organizzative.

Per affrontare il problema dei no-show e dell'uso subottimale delle agende, molti lavori adottano approcci di **IA predittiva** e **prescrittiva** mirati al contesto ambulatoriale. Un filone consistente utilizza modelli di *machine learning supervisionato* per predire la probabilità di no-show di ciascun paziente, identificando in anticipo quelli a rischio di mancata presentazione. Tali previsioni consentono di attuare strategie proattive, come il *trriage delle prenotazioni* e l'*overbooking controllato*: ad esempio, contattare preventivamente i pazienti ad alto rischio per confermare o riprogrammare l'appuntamento, oppure fissare pazienti aggiuntivi in slot che altrimenti rimarrebbero inutilizzati. Diversi studi mostrano che un uso intelligente dell'overbooking (*basato sull'IA*) permette di compensare le assenze senza incrementare i tempi d'attesa per gli altri pazienti. In alcuni casi, l'IA viene integrata nei sistemi informativi clinici in forma di dashboard interattive in tempo reale, attraverso cui i coordinatori possono monitorare i flussi di accesso e riassegnare rapidamente slot e personale in base alle necessità contingenti. L'adozione di queste soluzioni trasforma un approccio meramente predittivo in un

modello operativo più proattivo e adattivo, con l'obiettivo di ottimizzare la capacità ambulatoriale giorno per giorno.

Oltre alla previsione dei no-show, l'IA in ambito ambulatoriale viene impiegata per *prioritizzare e pianificare gli appuntamenti* in modo più efficiente rispetto al tradizionale ordine di arrivo. Ad esempio, Yousefi et al. (2020) propongono di sostituire i criteri di prenotazione "*first-come, first-served*" con una logica basata sul profilo clinico e la priorità del paziente. Nel loro studio, pazienti con maggiore urgenza clinica vengono individuati tramite algoritmi di clustering e ricevono slot anticipati, mentre un algoritmo di scheduling ottimizza la distribuzione degli appuntamenti tenendo conto delle capacità limitate e variabili del centro. Questo approccio di *triage data-driven* interpreta l'ambulatorio come un ecosistema dinamico in cui la gestione delle risorse è orientata al bisogno clinico, contribuendo a ridurre le disuguaglianze di accesso alle cure. In linea con i principi della sanità territoriale moderna, la pianificazione diventa così personalizzata e proattiva, privilegiando i pazienti più fragili e distribuendo il carico in modo più equilibrato.

Numerosi studi riportano miglioramenti tangibili grazie a queste soluzioni IA in ambito outpatient. Ad esempio, l'implementazione di un modello predittivo di no-show integrato nel workflow organizzativo di centri di cure primarie negli Emirati Arabi Uniti (EAU) ha portato a una riduzione del tasso di mancata presentazione dal 20,8% al 10,3% (*riduzione relativa ~50%*). Contestualmente, i tempi medi di attesa dei pazienti si sono abbassati di circa 5,7 minuti per visita, traducendosi in oltre 6.400 ore di attesa risparmiate nell'arco di tre mesi. Anche altre iniziative, come il *sistema di scheduling adattivo* descritto da Yousefi et al. (2020), hanno dimostrato di ridurre sensibilmente i tempi di attesa per i pazienti ad alta priorità e di aumentare il tasso di utilizzo delle risorse ambulatoriali disponibili. In quest'ultimo caso, l'integrazione tra modelli di *machine learning* (ML) e metodi di *operations research* (*processi di Markov decisionali*) ha consentito di coniugare efficienza e giustizia allocativa, due dimensioni spesso considerate in conflitto: il modello infatti riduce i ritardi e massimizza l'uso delle sale senza sacrificare l'equità di accesso, grazie a criteri di assegnazione più trasparenti e oggettivi. Un altro studio pionieristico (Lin et al., 2017) ha affrontato la *capacity planning* in un ambulatorio di oftalmologia tramite simulazione e ottimizzazione: il loro modello a due stadi (*allocazione risorse*

+ *block scheduling adattivo*) ha migliorato le performance del 53% rispetto allo scenario base *First-Come First-Served*, riducendo drasticamente sia i tempi di attesa dei pazienti sia le ore di straordinario del personale, il tutto senza aggiungere risorse rispetto all'organico esistente. Ciò dimostra il potenziale delle tecniche di *simulation-based optimization* nell'affinare la gestione di un ambulatorio complesso, bilanciando carichi di lavoro e comfort organizzativo attraverso l'uso di regole adattive di priorità e la riallocazione dinamica dello staff.

Va infine sottolineato che l'efficacia reale di questi strumenti in ambito ambulatoriale dipende non solo dalla loro accuratezza tecnica, ma anche dall'accettazione da parte di utenti e operatori. Uno studio qualitativo condotto nel NHS britannico (Wood et al., 2025) ha esplorato proprio la percezione e l'accettabilità di un sistema di gestione appuntamenti basato su ML. Dalle interviste a pazienti e staff sono emersi fattori chiave come la fiducia nella tecnologia, la trasparenza degli algoritmi e il coinvolgimento attivo degli operatori nel design del sistema. In generale, i pazienti hanno valutato positivamente strumenti come promemoria automatici e notifiche personalizzate (ritenendoli utili per non dimenticare gli appuntamenti e dunque migliorare l'aderenza).

Dal lato degli operatori sanitari, è stata enfatizzata la necessità che l'IA sia usabile e ben integrata nei flussi di lavoro, senza aumentare il carico ma anzi liberando tempo da dedicare alla cura. Questo indica che, oltre alle metriche classiche di efficienza, in ambiente ambulatoriale l'esito "soft" dell'accettabilità e della soddisfazione degli utenti finali rappresenta un outcome cruciale: la *fiducia* e il *buy-in* dello staff sono condizioni abilitanti perché i sistemi predittivi esprimano tutto il loro potenziale. In definitiva, nei contesti outpatient l'IA si sta affermando come leva per una gestione più proattiva, personalizzata e partecipativa delle agende, con benefici dimostrati su tempi di attesa, continuità assistenziale e utilizzo delle risorse – a patto di accompagnare l'innovazione tecnologica con strategie di change management centrate sulle persone.

4.1.2 Chirurgico/OR (stima durate, block scheduling, turnover, cancellazioni)

Un secondo grande ambito di applicazione riguarda il contesto chirurgico, in particolare la gestione delle sale operatorie (*Operating Rooms, OR*) e dei percorsi

perioperatori. Gli studi in questo settore riconoscono che il blocco operatorio è un ambiente ad alta complessità logistica e alto impatto economico, dove è necessario garantire la massima efficienza senza compromettere la sicurezza e la qualità delle cure. Tradizionalmente, la programmazione chirurgica si basava sull'esperienza dei clinici e su stime spesso conservative dei tempi operatori, comportando margini di sicurezza ampi ma anche periodi di inutilizzo delle sale. I risultati della letteratura recente mostrano come l'Intelligenza Artificiale stia trasformando questi processi, fornendo strumenti più rigorosi per prevedere e ottimizzare l'uso delle sale operatorie. In particolare, l'IA viene utilizzata per stimare con maggiore accuratezza la durata degli interventi, pianificare i blocchi operatori (*block scheduling*), gestire i tempi di turnover tra un caso e l'altro e anticipare possibili cancellazioni o ritardi.

Queste funzioni sono cruciali: una stima imprecisa dei tempi chirurgici può portare a lunghe attese e sforamenti di turno, mentre una stima troppo prudente può lasciare sale vuote; analogamente, le cancellazioni dell'ultimo minuto o un *turnaround* (tempo di riassetto e preparazione della sala operatoria tra due interventi) inefficiente tra interventi riducono il *throughput operatorio*. Di conseguenza, molti studi si concentrano su come l'IA possa razionalizzare l'utilizzo del blocco operatorio, bilanciando produttività e qualità dell'assistenza.

Gli approcci IA in ambito chirurgico sono prevalentemente basati su **modelli predittivi supervisionati** applicati a grandi moli di dati storici ospedalieri. Ad esempio, Martinez et al. (2021) hanno sviluppato un modello ML per la previsione della durata chirurgica a partire da oltre 80.000 interventi di 25 specialità diverse. Confrontando diversi algoritmi di apprendimento supervisionato – tra cui la **regressione lineare**, le **macchine a vettori di supporto** (*Support Vector Machine – SVM*), gli **alberi di regressione** e i **modelli ensemble basati su alberi** (*Bagged Trees*) – gli autori hanno riscontrato che i metodi tree-based ensemble offrivano la migliore accuratezza predittiva, con un errore quadratico medio (*Root Mean Squared Error, RMSE*) di circa 26 minuti.

Questo modello ha superato le performance del metodo tradizionale basato sull'esperienza dei programmatori chirurgici, riducendo l'errore medio di stima dei tempi operatori tra ~6 e 36 minuti in molte specialità comuni. Un risultato interessante è che, mentre l'umano tendeva a sovrastimare i tempi (*per sicurezza*), il

modello ML tendeva a sottostimarli leggermente. Ciò implica differenti tipi di bias: la sottostima algoritmica può generare ritardi a catena nella giornata operatoria, mentre l'eccesso di prudenza umano riduce l'utilizzo effettivo delle sale. Gli autori suggeriscono quindi di integrare meccanismi di calibrazione adattiva per allineare le previsioni alle esigenze organizzative specifiche.

Nel complesso, studi come quello di Martinez et al. dimostrano che algoritmi basati su alberi decisionali ed ensemble (ad es. *Random Forest*, *Gradient Boosting*) sono particolarmente efficaci nel predire variabili continue come la durata degli interventi in contesti reali caratterizzati da dati eterogenei e rumorosi. Una previsione più accurata dei tempi operatori si traduce in una saturazione più efficiente delle sale, meno slot vuoti e una progressione più regolare degli interventi durante la giornata. Di fatto, ridurre l'errore di stima contribuisce ad aumentare l'efficienza operativa e può alleviare lo stress sia per il personale sia per i pazienti (minori ritardi e attese).

Oltre ai tempi chirurgici, altre applicazioni predittive in sala operatoria riguardano la **gestione del flusso perioperatorio**. Diversi studi hanno sviluppato modelli per prevedere la permanenza in *recovery room*, ovvero nella *Post-Anesthesia Care Unit* (PACU), dopo l'intervento, oppure per identificare in anticipo i casi ad alto rischio di cancellazione chirurgica nel giorno programmato.

Ad esempio, Zhang et al. (2021) e Luo et al. (2020) hanno mostrato che algoritmi di Random Forest addestrati su dati clinici e logistici possono individuare con buona accuratezza ($AUC > 0,7$) i pazienti che probabilmente vedranno il loro intervento cancellato. Ciò consente al management di attuare misure preventive (come riempire il vuoto con altri casi o riprogrammare per tempo) per mantenere elevato il throughput operatorio.

Altri ricercatori si sono concentrati sul post-operatorio: modelli ML come quelli proposti da Mišić et al. (2020) e Hopkins et al. (2019) sono stati in grado di prevedere fino al 60% delle riammissioni entro 30 giorni dall'intervento, contribuendo a una gestione più personalizzata del rischio e potenzialmente a ridurre complicanze e costi ospedalieri. Queste applicazioni vanno oltre la pura programmazione chirurgica, estendendo l'uso dell'IA all'intero percorso assistenziale chirurgico, dalla preparazione alla dimissione.

Un filone di studi particolarmente rilevante in ambito chirurgico combina la predizione con l'ottimizzazione delle agende operatorie. In altre parole, l'output dei modelli predittivi viene utilizzato come input per algoritmi che elaborano la programmazione ottimale degli interventi, cercando di massimizzare l'utilizzo delle sale e minimizzare i tempi morti. Ad esempio, alcune ricerche integrano i tempi stimati degli interventi in modelli matematici di scheduling (*come la programmazione lineare intera, o euristiche avanzate*) per ottenere sequenze operative più aderenti alla realtà. Una revisione sistematica (Bellini et al., 2024) evidenzia che l'uso combinato di modelli predittivi ML con tecniche di ottimizzazione ha permesso di migliorare significativamente la puntualità dell'inizio degli interventi e di ridurre i tempi di inattività delle sale operatorie fino al 20%. In uno studio citato dagli autori, ad esempio, Lam et al. (2022) ha sviluppato modelli specifici per singolo chirurgo, migliorando del 18% l'accuratezza della previsione dei tempi e consentendo una pianificazione più realistica dei turni. Strömblad et al. (2021), in un trial clinico randomizzato, hanno dimostrato che l'uso di algoritmi ML nel scheduling riduce i tempi di attesa dei pazienti chirurgici senza diminuire la produttività complessiva del reparto.

Un altro esempio di approccio integrato è fornito da Rozario et al. (2020): implementando algoritmi ML all'interno del solver di ottimizzazione *Google OR-Tools*, il loro sistema di *intelligent scheduling* ha ottenuto in tre anni un risparmio economico di circa 469.000 dollari e una riduzione del 21% delle ore di straordinario infermieristico nel blocco operatorio. Questo risultato, oltre a evidenziare vantaggi in termini di costo e gestione del personale, sottolinea la sinergia tra IA e Ricerca Operativa: le previsioni fornite dall'IA diventano azioni concrete attraverso algoritmi di schedulazione automatizzata, massimizzando l'efficacia organizzativa.

Complessivamente, gli studi in ambito chirurgico riportano miglioramenti notevoli di performance grazie alle soluzioni di IA. In termini di outcome di efficienza, molte pubblicazioni documentano aumenti nell'utilizzo medio delle sale operatorie e riduzioni dei tempi morti intraoperatori. Nella revisione di Bellini et al., gli algoritmi analizzati hanno migliorato la precisione delle previsioni di durata mediamente del 10–30%, traducendosi in una diminuzione significativa dei tempi di inattività delle sale. Al contempo si registrano effetti positivi anche su metriche di qualità e

sicurezza: una programmazione più accurata e data-driven riduce il rischio di sovraccarico di casi (evitando ad esempio accavallamenti o congestione in recovery room) e migliora la gestione del personale sanitario, con carichi di lavoro più equilibrati. Diversi studi indicano anche una contrazione delle liste d'attesa chirurgiche e delle cancellazioni dell'ultimo minuto, segno di una maggiore stabilità e affidabilità delle agende. Ad esempio, Bellini et al. (2022) riportano che l'adozione diffusa di modelli ML in chirurgia ha consentito incrementi medi del *throughput operatorio* senza sacrificare la qualità percepita.

Parallelamente, Espaillat (2024) sottolinea come l'integrazione dell'IA nel contesto operatorio stia portando a una trasformazione sistemica più che tecnologica: per concretizzare il potenziale dell'IA in sala operatoria sono necessarie una solida governance dei dati, infrastrutture informatiche interoperabili con i sistemi ospedalieri e la formazione di competenze interdisciplinari (*clinici e data scientist*) che collaborino fianco a fianco.

Non a caso, alcune criticità ricorrenti emerse dai lavori includono la mancanza di validazioni esterne dei modelli, la variabilità dei dataset locali utilizzati e la limitata trasparenza di alcuni algoritmi complessi, tutti fattori che possono frenare l'implementazione pratica su larga scala. Gli esperti raccomandano pertanto l'adozione di soluzioni di *Explainable AI* e la standardizzazione di metriche organizzative, così da aumentare la fiducia nei sistemi di IA e permetterne il confronto oggettivo tra diversi studi e contesti.

In sintesi, nell'ambiente chirurgico l'Intelligenza Artificiale sta emergendo come uno strumento maturo e concreto per la razionalizzazione delle risorse operatorie. Essa consente un nuovo equilibrio tra produttività (*più interventi svolti nei tempi previsti*) e qualità assistenziale (*meno attese, minori rischi di complicanze da ritardi o affaticamento del team*). I benefici documentati – riduzione dei tempi di turnover, aumento dell'utilizzo delle sale, calo dei costi operativi e degli straordinari, minori cancellazioni – indicano che le soluzioni IA possono affrontare con successo alcune delle annose inefficienze dei blocchi operatori. Tuttavia, la letteratura mette anche in guardia sulla necessità di integrare queste tecnologie nei processi gestionali in modo strutturato, assicurando trasparenza, validazione e formazione: solo così sarà

possibile ottenere non solo algoritmi accurati, ma veri sistemi di supporto decisionale operativi accettati dal personale e sostenibili nel lungo periodo.

4.2 Per tecnologia IA

4.2.1 ML supervisionato (gradient boosting, random forest, prediction di no show e durate)

La stragrande maggioranza degli studi esaminati impiega tecniche di Machine Learning (ML) supervisionato, confermando come queste costituiscano il nucleo metodologico principale nelle applicazioni di IA per scheduling sanitario. In particolare, algoritmi tree-based e di ensemble learning emergono come strumenti altamente efficaci sia per la previsione dei no-show ambulatoriali sia per la stima delle durate chirurgiche. Tra questi, i modelli di Random Forest e Gradient Boosting (es. *XGBoost*, *CatBoost*) risultano tra i più utilizzati e performanti. Ad esempio, la revisione di Bellini et al. (2022) evidenzia una “netta prevalenza” di algoritmi ensemble come XGBoost, Random Forest e CatBoost nei sistemi di previsione per la sala operatoria.

In uno studio citato nella medesima revisione, XGBoost è emerso come il modello con le migliori prestazioni nella stima dei tempi chirurgici, superando sia le tecniche classiche (p.es. *regressione lineare*) sia approcci basati su analisi di serie storiche.

Analoghi riscontri si hanno in ambito ambulatoriale: diversi modelli di *gradient boosting* (CatBoost in particolare) hanno mostrato ottima accuratezza nell’identificare i pazienti a rischio di mancata presentazione, con valori di AUC compresi tra 0,64 e 0,87 a seconda dello scenario.

Accanto agli ensemble, continuano a essere impiegati – spesso come baseline o in modalità ibrida – anche algoritmi più tradizionali come la regressione logistica o gli alberi decisionali singoli. Tuttavia, vi è una chiara tendenza evolutiva: molte ricerche passano dalle semplici regressioni a modelli ML più complessi, in grado di catturare pattern non lineari e interazioni tra variabili. L’adozione crescente di tecniche come *random forest* e *gradient boosting* è motivata dalla loro capacità di gestire dati eterogenei, con variabili sia numeriche sia categoriali, e dalla robustezza rispetto

all'*overfitting* grazie al *bagging/boosting*. Ad esempio, come discusso (Martinez et al., 2021), un modello *Bagged Trees* ha ottenuto la migliore accuratezza globale nella previsione delle durate operatorie, riducendo l'errore rispetto ad altri algoritmi e al metodo umano. Analogamente, in contesti di no-show ambulatoriali, modelli basati su Random Forest riescono a identificare combinazioni di fattori demografici e clinici predittive dell'assenza con prestazioni superiori ai modelli lineari.

Un aspetto interessante emerso è l'attenzione alla interpretabilità anche nei modelli supervisionati complessi. Diversi studi affiancano agli algoritmi di ML strumenti di *Explainable AI (XAI)*, come l'*analisi delle feature importance* o i *Shapley values*, per comprendere quali variabili influenzino maggiormente le previsioni. Ciò riflette la necessità, in ambito sanitario, di bilanciare accuratezza predittiva e trasparenza: ad esempio, sapere che fattori come l'età del paziente, la specialità o l'ora del giorno contribuiscono alle previsioni di durata operatoria aiuta i clinici ad accettare e fidarsi del modello. Questa tendenza verso modelli più trasparenti e interpretabili è esplicitata in diversi lavori, dove agli algoritmi ensemble si accompagnano analisi post-hoc che spiegano le decisioni del modello. Nell'ambito del ML supervisionato applicato allo scheduling sanitario, si osserva una maturazione verso l'uso di metodi ensemble avanzati (*RF, boosting*) per massimizzare la performance predittiva, con un occhio alla spiegabilità e alla generalizzabilità (spesso perseguita tramite tecniche come il *cross-validation*, il *tuning di iperparametri* e, più di recente, approcci di *transfer learning* e *data augmentation* per arricchire i dataset disponibili).

Dal punto di vista dei risultati concreti ottenuti con i modelli supervisionati, i benefici sono molteplici: riduzione degli errori di previsione (*es. tempi operatori stimati con scarto medio ridotto, pazienti no-show individuati con alta sensibilità*), maggiore puntualità nelle schedule, e conseguentemente indicatori organizzativi migliori (*miglior utilizzo di sale e ambulatori, meno tempi morti, meno sorprese dell'ultimo minuto*). Inoltre, alcuni studi riportano che l'uso di modelli supervisionati ha permesso di scoprire pattern nascosti nei dati: ad esempio, individuare fasce orarie o combinazioni di procedura-equipe particolarmente critiche per ritardi, oppure riconoscere segmenti di popolazione di pazienti più inclini a disdire l'appuntamento. Queste informazioni estrapolate dal modello possono guidare interventi organizzativi mirati (come modificare l'allocazione delle risorse in certe

giornate o intensificare i promemoria per specifici pazienti). In definitiva, gli algoritmi di ML supervisionato – specialmente nelle forme evolute di ensemble (*Random Forest, XGBoost, CatBoost*) – rappresentano la spina dorsale delle soluzioni di IA in ambito ambulatoriale e chirurgico, offrendo predizioni accurate e affidabili che costituiscono il prerequisito per qualunque ottimizzazione successiva del sistema.

4.2.2 Deep Learning (RNN/LSTM, CNN – serie temporali di sala/attese)

L'utilizzo di tecniche di **Deep Learning** negli studi analizzati risulta meno frequente rispetto al ML "classico", ma rappresenta un settore di crescente interesse, soprattutto quando si tratta di gestire dati sequenziali, immagini o grandi moli di informazioni eterogenee. In ambito di gestione operativa di sale e ambulatori, il deep learning viene impiegato in scenari specifici dove occorre modellare complessi pattern temporali o combinare input multimodali. Ad esempio, reti neurali ricorrenti come i **modelli LSTM** (*Long Short-Term Memory*) sono state sperimentate per analizzare serie temporali relative a flussi di pazienti o parametri clinici nel perioperatorio. In uno studio riguardante il follow-up post- chirurgico, un modello LSTM addestrato su dati clinici longitudinali dei primi 12 mesi post-intervento ha raggiunto un'AUC media di ~0,90 nel predire esiti funzionali a lungo termine, superando nettamente altri algoritmi supervisionati tradizionali comparati nello studio. Ciò evidenzia la capacità del deep learning di catturare le dipendenze temporali e i trend di lungo periodo meglio di modelli statici, rendendolo adatto per prevedere l'andamento di variabili come il recupero del paziente o l'evoluzione di liste d'attesa su base temporale.

Nel contesto delle sale operatorie, il deep learning è stato impiegato soprattutto quando i dati a disposizione includono componenti non tabulari. Ad esempio, alcune applicazioni citate da Espaillat (2024) riguardano l'uso di **reti neurali convoluzionali** (*CNN*) per analizzare flussi video o immagini chirurgiche in tempo reale, identificare fasi operatorie e fornire feedback intraoperatorio sull'andamento della procedura. Sebbene queste applicazioni (*segmentazione di immagini, riconoscimento di gesti chirurgici, classificazione di tessuti*) escano in parte dal focus della pianificazione di risorse, esse rientrano nel più ampio panorama dell'IA a supporto della chirurgia

cognitiva e potrebbero indirettamente contribuire all'ottimizzazione dei tempi (*ad es. indicando in anticipo il completamento di una fase chirurgica e aiutando a preparare la successiva*). In tema di serie temporali di sala/attese, reti neurali profonde sono state sperimentate anche per prevedere l'andamento dell'occupazione di posti letto o sale nel tempo: ad esempio, modelli sequenziali multivariati che combinano LSTM e variabili d'ingresso come l'orario, il giorno della settimana, il tipo di intervento, hanno mostrato la capacità di anticipare i picchi di occupazione in recovery room meglio delle analisi di serie storiche tradizionali (spesso limitate da ipotesi di linearità).

Un altro campo in cui il deep learning compare è la fusione di dati multimodali per scopi decisionali: come riportato in letteratura, stanno emergendo **foundation models multimodali** e **approcci tipo Transformer** capaci di integrare simultaneamente immagini, testo e segnali fisiologici. Ad oggi, questi approcci avanzati sono più diffusi in contesti clinici diagnostici (*es. radiologia, patologia digitale*) che non nella programmazione di sale/operatori. Tuttavia, non mancano esempi in cui visione artificiale e NLP (*Natural Language Processing*) supportano la gestione operativa: un modello citato utilizza una CNN 3D combinata con un LSTM bidirezionale per analizzare automaticamente i video operatori e stimare il tempo rimanente di intervento, fornendo in output una predizione continua della durata residua della procedura. Ciò permette al team di sala di organizzarsi in anticipo per la fase di chiusura e per l'ingresso del caso successivo, riducendo i tempi morti. Analogamente, strumenti di NLP vengono impiegati per estrarre informazioni rilevanti da testi clinici non strutturati (*ad es. referti operatori, note anestesilogiche*) che possano influenzare la pianificazione – ad esempio identificando automaticamente dalle note i casi complessi che potrebbero sfiorare i tempi previsti.

In generale, l'apporto del deep learning nelle applicazioni di **scheduling** e **resource management sanitario** è ancora complementare a quello dei modelli supervisionati più semplici. Spesso le reti neurali profonde vengono affiancate ai metodi classici per vedere se aggiungono valore predittivo. Dai risultati emerge che in compiti ricchi di dati e complessità (*per es. previsione di complicanze con centinaia di variabili, analisi di segnali vitali in tempo reale*), il deep learning offre un boost di

performance (AUC fino a 0,94 in alcuni casi) grazie alla sua capacità di astrarre feature latenti dai dati. Di contro, richiede set di dati molto ampi per generalizzare bene e soffre di un'interpretabilità limitata, cosa che gli studi mitigano con tecniche XAI e validazioni prospettiche. Diversi studi prevedono un uso sempre più esteso di **reti neurali ricorrenti (RNN)** e di **reti neurali profonde (deep learning)**, progressivamente, cioè man mano che i dataset ospedalieri diventano più completi e integrati grazie all'aggiunta di nuove fonti di dati (*sensori, dispositivi IoT, cartelle cliniche elettroniche, ecc.*).

In conclusione, le tecniche di Deep Learning stanno aprendo nuove possibilità – ad esempio modellare dinamicamente l'intero percorso chirurgico del paziente o fornire sistemi di supporto intraoperatorio avanzato – che integrano e ampliano le funzionalità dei tradizionali modelli predittivi usati per scheduling e gestione attese.

4.2.3 Reinforcement Learning - politiche di scheduling adattivo

Il **Reinforcement Learning (RL)** rappresenta una frontiera emergente nel panorama delle tecnologie IA applicate alla gestione operativa sanitaria. A differenza del ML supervisionato, orientato alla predizione, il RL introduce una dimensione prescrittiva e adattativa, mirata a identificare strategie ottimali attraverso meccanismi di apprendimento per prove ed errori e feedback (*ricompense*). Negli studi esaminati, i contributi che impiegano RL puro per problemi di scheduling sono ancora limitati, ma diversi autori ne sottolineano il potenziale. In generale, un algoritmo di RL può apprendere politiche di scheduling adattivo sperimentando diverse azioni (*ad es. diverse allocazioni di slot o priorità di chiamata*) e ricevendo ricompense in base all'efficienza ottenuta (*ad es. minima attesa residua, massima occupazione delle risorse*). Iterando questo processo, il sistema impara la sequenza di decisioni che massimizza la ricompensa cumulativa, ovvero l'esito organizzativo desiderato.

Un esempio concreto di applicazione concettuale del RL in sanità viene dalla gestione di terapie e complicanze: Loftus et al. (2020) descrivono come il reinforcement learning possa individuare strategie terapeutiche ottimali, mostrando in test controllati performance superiori a quelle umane nella gestione di scenari complessi (come la sepsi). Traslando questo paradigma nel dominio dello scheduling, si può immaginare un algoritmo di RL che autonomamente aggiusti le

regole di pianificazione in un ambulatorio o in un blocco operatorio, adattandosi dinamicamente alle condizioni operative reali. Ad esempio, un RL potrebbe apprendere quante visite aggiuntive inserire come **overbooking** in diverse fasce orarie per minimizzare sia i tempi morti sia il rischio di sovraccarico, calibrandosi sui risultati giornalieri. Oppure, in sala operatoria, potrebbe decidere in tempo reale l'ordine successivo dei casi da una lista di attesa, in base all'andamento effettivo della giornata chirurgica (anticipando interventi brevi se si accumula ritardo, ecc.), perseguendo la massima puntualità a fine giornata.

Nei lavori esaminati, alcune componenti di learning adattivo sono talvolta presenti anche senza citare esplicitamente il RL. Ad esempio, il **Markov Decision Process (MDP)** utilizzato da Yousefi et al. (2020) per la scheduling ambulatoriale può essere visto come una formulazione decisionale sequenziale (*stocastica*) che, se risolta tramite algoritmi di iterazione di politiche, si avvicina al paradigma del reinforcement learning. In altri studi, come la revisione di Entezari et al. (2023), si fa riferimento a modelli markoviani e simulazioni Monte Carlo per ottimizzare l'allocazione di risorse chirurgiche in modo dinamico. Inoltre, articoli di prospettiva sul futuro della chirurgia digitale vedono nel reinforcement learning un approccio chiave per l'ottimizzazione dinamica dei processi chirurgici, ad esempio adattando in tempo reale i parametri di una procedura robotica in risposta a ciò che accade in sala.

Sebbene dunque non si riscontrino ancora implementazioni estese di RL "puro" per la gestione di agende e risorse, la letteratura concorda sul fatto che questa tecnica offra un paradigma potentemente generativo di policy. Il RL potrebbe fungere da complemento ai modelli predittivi: laddove un modello ML prevede cosa accadrà (es. *"quanti pazienti non si presenteranno domani?"*), un agente RL può suggerire cosa fare (*"come riorganizzare gli appuntamenti di domani per massimizzare l'efficienza data la previsione di X no-show"*).

Un aspetto interessante è che il RL, avendo come obiettivo una ricompensa definita dall'utente, permette di incorporare in maniera flessibile diverse metriche: si può addestrare l'agente a massimizzare un indice composito che bilanci, ad esempio, tempi di attesa, tasso di occupazione e overtime del personale. In letteratura, gli autori sottolineano che questi approcci di RL necessitano però di grandi quantità di

dati e simulazioni fedeli dell'ambiente reale per addestrare efficacemente l'agente. Inoltre, come tutte le soluzioni **data-driven**, richiedono robusti meccanismi di validazione prima di un'adozione clinica, per evitare risultati indesiderati o policy non sicure.

In sintesi, il Reinforcement Learning rappresenta un approccio promettente per sviluppare sistemi di scheduling auto-adattativi, capaci di *“imparare dall'esperienza”* e di migliorare progressivamente le proprie decisioni. I pochi esempi emergenti (*spesso ancora a livello di simulazione o prototipo*) suggeriscono che il RL può contribuire a colmare il gap tra previsione e azione, fornendo raccomandazioni ottimali su come riorganizzare risorse e appuntamenti in risposta alle condizioni correnti. Sebbene ancora non diffuso nei 52 studi inclusi, il RL viene indicato dagli esperti come una naturale evoluzione dei sistemi di pianificazione: uno strumento in grado di andare oltre le regole statiche e adattarsi in tempo reale, avvicinandosi all'obiettivo di una *“intelligenza organizzativa”* autonoma per le strutture sanitarie.

4.2.4 Ibridi ML + Ricerca Operativa/Programmazione matematica (cluster predict & schedule, ottimizzazione vincolata)

Una tendenza trasversale emersa con forza dagli studi è l'approccio ibrido, che combina algoritmi di Machine Learning con metodi di Ricerca Operativa (RO) o Programmazione Matematica tradizionale. L'idea di fondo è sfruttare i punti di forza complementari di ciascun paradigma: i modelli ML eccellono nell'apprendere pattern dai dati (*es. prevedere durate, no-show, rischi*), mentre gli algoritmi di Ricerca Operativa sono ideali per risolvere problemi di ottimizzazione vincolata (*es. scheduling di slot, assegnazione di risorse*) una volta definiti parametri e obiettivi. Dall'unione nasce quella che alcuni autori chiamano una vera e propria **“Intelligenza Artificiale operativa”**, in cui la capacità predittiva si traduce direttamente in decisioni automatizzate ottimali.

Molti lavori recenti indicano che le strategie più efficaci per migliorare efficienza e outcome organizzativi derivano proprio dall'integrazione di modelli predittivi ML con modelli di ottimizzazione. Ad esempio, nella già citata revisione di Entezari et al. (2023) dedicata al percorso chirurgico di artroprotesi, la seconda parte del lavoro analizza studi di ottimizzazione (*mixed-integer programming, simulazioni Monte*

Carlo, modelli markoviani) applicati a problemi come l'allocazione dei blocchi operatori, la gestione dei letti post-operatori, la programmazione del personale e la riduzione dei tempi morti. Il messaggio chiave degli autori è che l'unione tra predizione e ottimizzazione è la chiave per una chirurgia più efficiente e sostenibile: i risultati complessivi mostrano miglioramenti significativi nell'utilizzo delle sale, riduzioni della degenza (*Length of Stay - LOS*) e agende chirurgiche più stabili proprio quando le tecniche di IA avanzata vengono combinate con modelli matematici decisionali.

Di converso, lavorare con approcci isolati (*solo predizione o solo ottimizzazione*) risulta meno incisivo. Ad esempio, Entezari menziona che la strategia a due stadi implementata in alcuni studi – prima un modello ML che prevede indicatori chiave come durata intervento e lunghezza degenza, poi un modello di ottimizzazione che usa queste previsioni per allocare risorse – supera le strategie singole e permette un balzo di performance specialmente in presenza di variabilità operativa. Questo approccio ibrido consente di anticipare i fabbisogni e di orchestrare l'intero percorso perioperatorio come un sistema dinamico e adattivo, piuttosto che statico e frammentato.

Un esempio concreto di architettura ibrida Machine Learning e Ricerca Operativa in ambito ambulatoriale è offerto dallo studio di Yousefi et al. (2020). Esso prevede una prima fase di clustering dei pazienti (*ML non supervisionato*) per assegnare classi di priorità clinica, seguita da una seconda fase di Markov Decision Process che calcola la politica ottimale di scheduling per i giorni successivi. In questo modo, la previsione (*quali pazienti avranno bisogno di accesso più tempestivo*) alimenta direttamente l'ottimizzazione (*come distribuire gli slot disponibili alle varie classi di priorità*), creando un flusso integrato che va dal dato grezzo alla decisione operativa. Gli autori sottolineano che tale approccio ibrido realizza un sistema di supporto decisionale intelligente, in cui l'IA non si limita a fare analisi ma genera automaticamente regole decisionali implementabili quotidianamente. Il risultato è una **pianificazione adattiva**, verificata nella pratica, che combina i vantaggi della flessibilità data-driven con quelli della robustezza matematica (*il MDP garantisce soluzioni vicine all'ottimo teorico ma computazionalmente gestibili grazie ad approssimazioni*). Anche in altri contesti gli ibridi ML+RO hanno dato prova di

efficacia: nel **block scheduling chirurgico**, ad esempio, l'uso di algoritmi di ottimizzazione vincolata (*come la mixed-integer programming*) alimentati da previsioni accurate di durata degli interventi ha consentito di minimizzare i tempi inutilizzati e di ridurre sia le attese sia gli straordinari. Similmente, per la gestione di ambulatori sovraffollati, approcci di **simulation- optimization** (come quello di Lin et al., 2017) hanno mostrato di poter sperimentare virtualmente diversi scenari di allocazione risorse e scheduling, identificando soluzioni ottimali che poi vengono applicate al mondo reale con sostanziali miglioramenti di performance.

Gli studi ibridi evidenziano un ulteriore beneficio: la possibilità di includere multi-obiettivo e considerazioni pratiche nel modello. Ad esempio, Lin et al. (2017) definiscono una funzione obiettivo che combina tre indicatori chiave – tempi di attesa, ore di straordinario, livello di congestione in sala d'attesa – ponderandoli secondo le priorità organizzative. L'algoritmo ricerca quindi una soluzione di compromesso che bilanci efficienza e qualità del servizio. All'occorrenza, mediante simulazione, è possibile testare varie impostazioni (*ad es. dare più peso alla riduzione attese vs riduzione straordinari*) e scegliere la configurazione ritenuta migliore dai decisori umani. Questa capacità di effettuare analisi "*what-if*" attraverso modelli ibridi è preziosa: come riportato da Entezari et al. (2023), alcune applicazioni di supporto decisionale incluse nella loro revisione usano simulazioni per valutare scenari alternativi e differenti politiche gestionali, mostrando come l'IA possa evolvere da strumento analitico a partner operativo per la pianificazione strategica. In altre parole, l'IA ibrida può fungere sia da progettista (*generando soluzioni ottimali*) sia da simulatore (*valutando le conseguenze di scelte possibili*), offrendo ai manager sanitari una visione a 360 gradi delle opzioni.

In conclusione, la stratificazione per tecnologia evidenzia che gli approcci ibridi ML + RO sono tra i più promettenti e incisivi nel migliorare la gestione di ambulatori e sale operatorie. Essi combinano predizione accurata e ottimizzazione efficiente, traducendo il "*sapere*" fornito dai dati nel "*fare*" di decisioni operative ottimali. Questa sinergia consente di ottenere risultati tangibili in termini di efficienza e sostenibilità organizzativa, come dimostrato da vari esempi (*riduzione del 50% delle attese in clinica, risparmi economici nell'ordine di centinaia di migliaia di euro in chirurgia*). La sfida futura sarà diffondere e standardizzare tali approcci, curando gli

aspetti di integrazione nei sistemi informativi sanitari e la formazione del personale, così da sfruttare a pieno quella che appare come la frontiera avanzata dell'IA applicata alla sanità operativa.

4.3 Per compito/uso

4.3.1 Previsione (durata interventi, no show, lunghezza lista)

La previsione (*prediction*) rappresenta la funzione fondamentale e più diffusa delle applicazioni di IA nei contesti analizzati. Predire eventi o variabili chiave – prima che si verifichino – consente ai decisori di pianificare meglio le risorse e di anticipare problemi. Nei 52 studi, le principali variabili target oggetto di previsione sono: i **tempi di durata degli interventi chirurgici**, la **probabilità di no-show** dei pazienti agli appuntamenti, e talvolta indicatori globali di volume come la **lunghezza di liste di attesa** o il **numero di pazienti in arrivo** (*domanda*).

Come emerso in precedenza, moltissimi lavori si concentrano sulla previsione accurata della durata operatoria di interventi chirurgici. Questo è un parametro critico: stimare in anticipo se un intervento durerà 60 minuti o 120 minuti fa la differenza tra una giornata chirurgica che scorre puntuale e una che accumula ritardi a cascata. I modelli di ML supervisionato (*alberi, ensemble*) hanno permesso di ridurre sensibilmente l'errore di stima rispetto ai metodi tradizionali basati sull'esperienza personale. Ad esempio, il modello di Martinez et al. (2021) con Bagged Trees ha raggiunto un RMSE (*Root Mean Squared Error*) di ~26 minuti sul set di test, contro errori ben superiori (anche >40 minuti) riscontrati con approcci più semplici. Ciò significa che, in media, l'IA riesce a prevedere i tempi con uno scarto di mezz'ora, dove prima l'errore poteva essere di un'ora o più. Anche altri studi citati (es. Zhao et al., 2019) riportano incrementi nell'accuratezza predittiva dei tempi operatori dell'ordine del +16-18% grazie all'analisi di dati storici con modelli ML. Naturalmente, la precisione non sarà mai assoluta per via della variabilità intrinseca in chirurgia, ma una riduzione dell'errore di alcuni minuti per caso si traduce – su decine di casi al mese – in ore guadagnate o ritardi evitati. Va anche notato che la previsione della durata spesso include la stima dei tempi di preparazione e turnover

(tra anestesia, posizionamento paziente, pulizia sala, ecc.), estendendo quindi la predizione oltre il puro “tempo di bisturi”. Alcuni modelli vengono addestrati specificamente su componenti del processo (ad es. durata dell’induzione anestesiologicala vs durata del taglio) per fornire un quadro più dettagliato ai fini organizzativi.

Sul versante ambulatoriale, il compito predittivo più diffuso è identificare in anticipo i pazienti che non si presenteranno all’appuntamento. Anche in questo caso, i modelli di ML (tipicamente classificazione binaria con random forest, boosting o reti neurali) hanno dimostrato buona efficacia. Ad esempio, AlSerkal et al. (2025) hanno sviluppato un modello di Random Forest che classifica i pazienti in tre categorie di rischio no-show (alto, medio, basso): il modello ha raggiunto un’accuratezza complessiva dell’86% nell’individuare i casi ad alto rischio. Ciò significa che la stragrande maggioranza dei pazienti che poi effettivamente non venivano all’appuntamento era stata correttamente segnalata dal sistema in anticipo. In letteratura si trovano anche metriche come l’AUC per questi modelli, spesso superiori a 0,80–0,85, indice di una buona capacità discriminante. Variabili predittive comuni includono: storico delle precedenti mancate presentazioni, anticipo di prenotazione (i booking fatti con molto preavviso hanno più rischi di no-show), caratteristiche demografiche e cliniche, giorno della settimana, promemoria inviati, ecc. La combinazione di più fattori tramite ML permette di superare approcci semplici (es. regole basate su singoli criteri come “i pazienti giovani mancano più spesso”, che sono spesso inaffidabili). La revisione di Toffaha et al. (2025) sottolinea proprio come si sia passati da modelli predittivi semplici a modelli ML più complessi e ibridi, capaci di apprendere pattern comportamentali dei pazienti e di adattarsi ai diversi contesti (es. diverse specialità ambulatoriali). Il risultato è che oggi disponiamo di strumenti predittivi che – se alimentati con dati di qualità – possono ridurre significativamente l’impatto dei no-show, prevedendoli e consentendo di agire di conseguenza.

Un altro oggetto di previsione rilevante è la lunghezza di lista o, più in generale, il volume di domanda attesa in un certo periodo. In alcuni studi ciò si traduce nel prevedere quanti pazienti arriveranno in pronto soccorso o in ambulatorio in un dato giorno, oppure quanti interventi chirurgici di una certa classe dovranno essere

programmati nel prossimo mese. Queste previsioni di “*workload*” servono per il **capacity planning strategico**. Ad esempio, modelli di forecasting statistico e ML possono essere applicati per predire il numero di procedure chirurgiche settimanali basandosi su trend storici e segnali esterni (*festività, stagionalità influenzali, etc.*), aiutando a definire quanti blocchi operatori aprire o quante sedute ambulatoriali aggiuntive pianificare.

Anche se non tutti gli studi includono esplicitamente questo tipo di previsione, alcuni – specie quelli focalizzati su ottimizzazione – lo considerano implicitamente: il modello di Lin et al. (2017), ad esempio, simula scenari con vari numeri di pazienti e vari livelli di no-show per testare la robustezza delle soluzioni di scheduling. In ambito di liste di attesa chirurgiche, l’IA può essere impiegata per stimare il **time-to-surgery** per un paziente appena inserito in lista, tenendo conto delle priorità cliniche e del backlog corrente; oppure per predire quanti pazienti in lista saranno operati entro un certo termine, dati gli slot disponibili e il tasso di nuovi ingressi. Queste informazioni permettono di identificare colli di bottiglia e di intervenire, ad esempio attivando sedute operatorie aggiuntive o programmando convenzioni con strutture esterne.

La funzione predittiva dell’IA fornisce il basamento informativo su cui poi costruire decisioni e interventi. Prevedere con ragionevole accuratezza cosa accadrà (*quanti pazienti arriveranno, chi non verrà, quanto durerà un intervento*) cambia il paradigma gestionale da reattivo a proattivo. I risultati documentati mostrano miglioramenti in vari indicatori: riduzione dell’errore medio di stima (*tempi più vicini a quelli reali*), aumento dell’adesione agli appuntamenti (*meno mancati accessi grazie alle azioni sui pazienti a rischio*), maggiore prevedibilità complessiva dei percorsi di cura (*meno variabilità inattesa*). Ciò getta le basi perché le strutture sanitarie possano poi ottimizzare e pianificare con efficacia (*un planning è tanto migliore quanto più affidabili sono le previsioni su cui si fonda*).

In definitiva, “**conoscere prima**” tramite i modelli predittivi permette di “**agire meglio dopo**”, e quasi tutti gli studi concordano su questo aspetto, ponendo la predizione al centro dei sistemi di clinical operations intelligence in sviluppo.

4.3.2 Ottimizzazione (allocazione slot/risorse, block management, riorganizzazione agende)

La seconda macro-funzione identificata negli studi è l'**ottimizzazione**: dato un certo obiettivo e certi vincoli, l'IA (*spesso in combinazione con metodi di ricerca operativa*) cerca di individuare la configurazione ottimale per l'allocazione di slot, risorse o appuntamenti. In altre parole, mentre la previsione risponde alla domanda "cosa accadrà se...?", l'ottimizzazione risponde a "qual è il modo migliore di...?". Nei contesti sanitari analizzati, i problemi di ottimizzazione più frequenti riguardano: l'allocazione degli slot di calendario ai pazienti (*schedulazione degli appuntamenti o degli interventi*), la gestione dei blocchi operativi (*es. quanti e quali interventi inserire in ciascuna sala, o come suddividere il tempo tra diverse specialità*), l'allocazione del personale e delle risorse (*turni, letti, sale diagnostiche*), e la riorganizzazione delle agende in seguito a eventi dinamici (*ad es. cancellazioni, urgenze sopravvenute, variazione di domanda*).

Diversi studi propongono modelli per l'ottimizzazione dell'overbooking ambulatoriale: una volta predetti i no-show, si tratta di decidere quanti e quali pazienti aggiuntivi inserire in agenda in modo da massimizzare l'utilizzo senza creare eccessiva attesa. Questo è un classico problema di bilanciamento rischio-beneficio, che può essere affrontato con modelli matematici (*p. es. ottimizzazione stocastica*) che massimizzano una funzione di utilità attesa. Alcuni lavori, come ricordato, hanno integrato modelli predittivi con euristiche di scheduling per definire politiche automatiche di prenotazione: ad esempio, Yousefi et al. (2020) derivano una **benchmark policy** che a fine di ogni giornata suggerisce quanti appuntamenti assegnare in ciascuna classe di priorità per i giorni successivi, minimizzando costi di ritardo e mancata assegnazione. Tale politica, ottenuta risolvendo in modo approssimato un MDP (*Markov Decision Process*), si è dimostrata vicina all'ottimo teorico ma molto più semplice da implementare operativamente. In altri studi, l'ottimizzazione dell'agenda si traduce nello scegliere l'ordine ottimale dei pazienti: ad esempio, ordinare le visite in modo da ridurre i tempi di attesa totali o da tenere conto delle diverse durate previste (*magari alternando visite lunghe e corte per evitare accumuli in coda*). Algoritmi di scheduling adattivo sono stati testati sia in ambito radiologico (*per minimizzare i tempi di attraversamento in centri diagnostici*

complessi) sia in ambito di visite specialistiche, con miglioramenti rispetto al *first-come-first-served* in termini di tempo medio trascorso in struttura per paziente.

Nella gestione del blocco operatorio, i compiti di ottimizzazione assumono un ruolo centrale. Qui troviamo problemi come il **block scheduling** (*stabilire l'assegnazione di blocchi di tempo in sala alle diverse specialità o chirurghi, tipicamente su base settimanale/mensile*) e il **case scheduling** (*decidere la sequenza e l'accoppiamento intervento-sala specifico per ogni giornata operativa*). Diversi studi citati nelle revisioni hanno utilizzato modelli di *Mixed-Integer Linear Programming (MILP)* per assegnare i blocchi operatori in modo da massimizzare l'utilizzo e minimizzare il rischio di overtime. Questi modelli tengono conto di vincoli come: disponibilità di chirurghi, priorità cliniche dei casi (*che magari impongono finestra massima di attesa*), compatibilità sala-attrezzature, turni anestesisti/infermieri, ecc. L'IA in sé entra spesso nel fornire input dinamici (*es. numero di casi attesi, durata stimata*) a questi modelli, che poi risolvono il problema combinatorio. Ad esempio, Bellini et al. (2019, citato in Espaillet 2024) mostrano come modelli di ottimizzazione integrati con previsioni di ritardi e cancellazioni tramite Random Forest possano coordinare simultaneamente sale operatorie e recovery room, ottenendo un flusso perioperatorio più fluido. Un altro aspetto dell'ottimizzazione in sala è la gestione dei tempi di turnover: alcuni lavori propongono soluzioni per allocare in modo ottimale le risorse di supporto (*infermieri di sala, tecnici*) durante i cambi di caso, oppure per decidere l'ordine dei casi tale da ridurre i tempi morti (*ad es. programmare prima un caso breve seguito da uno lungo per dare tempo alla sterilizzazione strumenti, etc.*). Simulazioni Monte Carlo sono state utilizzate per testare migliaia di possibili schedulazioni giornaliere e selezionare quella con minore tempo totale inattivo.

L'**ottimizzazione delle risorse umane** è un altro ambito toccato: per esempio, Lin et al. (2017) nel loro modello a due stadi dedicano la "*Stage I*" proprio all'allocazione ottimale del personale disponibile in ambito ambulatoriale, in base alla flessibilità di competenze. Il loro algoritmo iterativo rialloca dinamicamente infermieri e medici ai vari compiti, cercando la configurazione che minimizza attese e straordinari. Questo è particolarmente rilevante in tempi di carenza di personale: un'IA che suggerisca come distribuire al meglio il personale esistente tra le diverse attività (*ad es.*

bilanciando ambulatori e piccoli interventi) può contribuire a mantenere alti livelli di servizio senza nuove assunzioni. Alcuni studi inoltre considerano la **riorganizzazione reattiva delle agende**: se durante la giornata accade un evento (*un chirurgo si ammala, un paziente cancella all'ultimo momento, arriva un'urgenza*), come *"riorganizzare il resto della giornata minimizzando l'impatto?"*. Qui troviamo applicazioni di algoritmi di ripianificazione in tempo reale, talvolta basati su **Reinforcement Learning** o su **euristiche greedy** guidate da modelli predittivi. Ad esempio, si può ricalcolare la sequenza ottima dei restanti interventi quando uno slitta, oppure redistribuire i pazienti di un ambulatorio su altri medici se uno studio deve chiudere anticipatamente.

I benefici riportati grazie alle ottimizzazioni IA-driven sono notevoli. Riprendendo alcuni dati salienti: Lin et al. (2017), hanno ottenuto un miglioramento medio del 53% rispetto allo scenario base, con riduzioni marcate di tempi di attesa e straordinari grazie alla strategia ottimizzata. Rozario et al. (2020), hanno raggiunto risparmi economici e di ore-personale significativi migliorando la programmazione operatoria. In generale, molti studi riferiscono riduzioni dei tempi morti (*sia ambulatoriali che intraoperatori*) dell'ordine del 10-20%, e incrementi di produttività (*numero di pazienti visti o interventi eseguiti per unità di tempo*) altrettanto rilevanti. Va sottolineato che spesso l'ottimizzazione raggiunge questi risultati senza investimenti aggiuntivi in risorse, ma semplicemente utilizzando meglio quelle esistenti. Si tratta quindi di guadagni di efficienza pura, che in ambito sanitario si traducono in più prestazioni erogate o riduzione di costi operativi, a parità di struttura.

La dimensione del compito ottimizzativo vede l'IA operare come un **"motore decisionale"** che suggerisce come organizzare appuntamenti, sale e personale per ottenere i migliori risultati possibili. Se la fase predittiva fornisce gli elementi (*durata attesa di X, rischio Y*), la fase di ottimizzazione li combina per fornire un piano d'azione (*agenda, orario, assegnazione*) che massimizza l'efficienza o altri criteri target. Dalla letteratura emerge chiaramente che le soluzioni più avanzate integrano strettamente predizione e ottimizzazione (*cf. capitolo 4.2.4*), e che questa integrazione porta i maggiori benefici. L'IA consente di affrontare problemi di scheduling complessi, con molti vincoli e obiettivi, trovando soluzioni che

difficilmente un pianificatore umano potrebbe individuare manualmente. Con l'aumento della capacità computazionale e la disponibilità di dati in tempo reale, si prospetta un futuro in cui le agende sanitarie saranno continuamente ricalcolate e ottimizzate in risposta allo stato del sistema, attraverso algoritmi adattivi – in linea con il concetto di scheduling intelligente e dinamico.

4.3.3 Supporto decisionale operativo (Operational Decision Support Systems – ODSS): simulazioni “what-if” e dashboard

Oltre a predire e ottimizzare i processi, molte applicazioni di intelligenza artificiale in sanità operano anche come strumenti di **supporto decisionale operativo**, rientrando nella categoria degli **Operational Decision Support Systems (ODSS)**. Questi sistemi forniscono a clinici e manager strumenti utili per comprendere scenari complessi e prendere decisioni informate. In questa categoria rientrano le simulazioni di scenario (analisi “what-if”) e le dashboard interattive basate su tecniche di analisi avanzata dei dati, che consentono di visualizzare l'andamento dei processi e di sperimentare virtualmente le conseguenze di diverse scelte organizzative. Diversi studi adottano un approccio di **simulation-optimization**, costruendo un modello simulativo del processo sanitario (ambulatorio o percorso chirurgico) e utilizzandolo come banco di prova per testare differenti politiche organizzative. Ad esempio, la revisione di Entezari et al. (2023) menziona studi che utilizzano simulazioni **Monte Carlo** per valutare diverse strategie di allocazione dei letti post-operatori o di programmazione chirurgica, confrontando metriche quali il tasso di utilizzo delle risorse e i costi associati a ciascuno scenario. In questo modo, le simulazioni diventano strumenti particolarmente efficaci di **supporto alle decisioni operative**: i dirigenti possono porre domande del tipo “*cosa succederebbe se aumentassimo del 10% gli interventi ortopedici il prossimo mese?*” oppure “*che impatto avrebbe sulla lista d'attesa ridurre da 30 a 20 giorni il tempo medio tra prenotazione e intervento per i casi oncologici?*” e ottenere risposte quantitative dal modello. Una caratteristica chiave è che l'IA consente di costruire simulazioni più realistiche, perché i modelli predittivi incorporati forniscono parametri accurati (*es: distribuzione delle durate, probabilità di complicanze, tassi di no-show*) che alimentano la simulazione. Il risultato è una valutazione affidabile di scenari ipotetici senza doverli provare sul campo, riducendo i rischi di decisioni al buio.

Un esempio concreto di supporto decisionale tramite simulazione è lo studio di Lin et al. (2017): oltre a ottimizzare, il loro sistema consente di variare i pesi degli obiettivi (*attesa vs straordinari vs congestione*) per vedere come cambia la soluzione ottima, fornendo così insight su quali trade-off l'organizzazione è disposta ad accettare. Allo stesso modo, altri studi hanno creato modelli digitali di ospedale (*digital twin*) in cui testare, ad esempio, l'effetto di introdurre una nuova sala operatoria o di modificare gli orari ambulatoriali, prima di attuare realmente tali modifiche.

Parallelamente, il supporto decisionale in tempo reale si manifesta tramite dashboard e sistemi di allerta. Uno dei casi più interessanti è il progetto di AISerkal et al. (2025): qui un modello predittivo di no-show è stato integrato in un cruscotto operativo collegato all'**Electronic Health Record - EHR** (*ovvero, la Cartella Clinica Elettronica - CCE*), visibile ai coordinatori di centri di assistenza primaria. La dashboard mostrava in tempo reale gli appuntamenti del giorno con un indicatore di rischio di no-show accanto a ciascun paziente, permettendo allo staff di intervenire prontamente (*telefonate di recall, rimpiazzo di pazienti*). Inoltre visualizzava metriche aggregate (*ad es. tasso di no-show attuale vs storico, tempi medi di attesa aggiornati*) e consentiva la riallocazione immediata del personale tra diverse aree in base ai picchi di affluenza, grazie al monitoraggio continuo dei flussi. Questo sistema ha trasformato l'attività di gestione quotidiana in una pratica data-driven: i decision maker dispongono di informazioni aggiornate e predittive per aggiustare il piano operativo "*al volo*". I risultati, come citato, sono stati notevoli (*dimezzamento dei no-show e riduzione significativa dei tempi di attesa*), a testimonianza che una buona informazione in tempo reale può migliorare gli outcome tanto quanto (*o in sinergia con*) complessi algoritmi offline.

Altre forme di supporto decisionale includono i sistemi di allerta precoce e le raccomandazioni automatizzate. Ad esempio, in ambito chirurgico, piattaforme come **MySurgeryRisk** (*Università della Florida*) integrano modelli predittivi di complicanze postoperatorie e forniscono al chirurgo e all'anestesista un'interfaccia decisionale che segnala pazienti ad alto rischio e suggerisce possibili azioni (*monitoraggi aggiuntivi, posti in terapia intensiva riservati, ecc.*). Sebbene questo sia più clinical decision support che operations, impatta la gestione delle risorse

postoperatorie (*letti di Terapia Intensiva, tempi di recupero*) e rientra nel concetto di IA a supporto delle decisioni in tutto il percorso assistenziale. Un altro esempio: sistemi che analizzano in tempo reale i flussi di pazienti al triage (*magari con telecamere o sensori IoT in sala d'attesa*) potrebbero allertare i responsabili se si sta formando una coda troppo lunga, raccomandando di deviare personale da altre attività per supportare il front desk. In letteratura si inizia a parlare di *ambient intelligence* e *sensor-enhanced OR* dove l'IA percepisce l'ambiente e offre suggerimenti contestuali al personale.

Riassumendo, la dimensione del supporto decisionale copre tutti quegli usi dell'IA in cui l'obiettivo non è sostituire il decisore, ma mettergli a disposizione informazioni e strumenti per decidere meglio. Le simulazioni "*what-if*" aiutano nella pianificazione strategica e tattica, riducendo l'incertezza delle scelte a medio-lungo termine. Le dashboard e gli alert basati su AI aiutano nella gestione operativa giornaliera e nell'immediato, mantenendo il sistema flessibile e adattabile agli imprevisti. Entrambi convergono verso l'idea di un'intelligenza aumentata: il decision maker umano resta al centro, ma è potenziato da analisi complesse e continue che l'IA svolge dietro le quinte. Gli studi dimostrano che questo porta a sistemi sanitari più reattivi e proattivi al tempo stesso – *capaci di simulare il futuro e reagire in tempo reale* – con benefici su efficienza e qualità. Non a caso, nella revisione di Toffaha et al. (2025), si sottolinea che l'IA viene usata non solo per predire no-show ma anche come strumento di supporto per ottimizzare strategie di overbooking, organizzazione delle agende e uso efficiente del personale. Il trait d'union di queste attività è proprio il supporto decisionale fornito all'uomo, che può così prendere decisioni basate su evidenze data-driven piuttosto che su intuizioni o regole fisse.

4.4 Per outcome

4.4.1 Efficienza (tempi di attesa, utilizzo sale/ambulatori, throughput, turnover, cancellazioni)

Gran parte degli studi esaminati valuta il successo delle soluzioni di IA attraverso metriche di efficienza organizzativa. Migliorare l'efficienza significa in sostanza fare un uso migliore del tempo e delle risorse disponibili, riducendo sprechi e colli di

bottiglia. Tra gli indicatori più frequentemente monitorati troviamo: i **tempi di attesa** (*dei pazienti per ottenere la prestazione, o attesa in sede*), il **tasso di utilizzo di sale operatorie e ambulatori**, il **throughput** (*numero di pazienti trattati in un certo periodo*), i **tempi di turnover in sala operatoria**, e il **numero di cancellazioni o slot rimasti vuoti**.

I risultati raccolti indicano miglioramenti significativi su tutti questi fronti grazie all'implementazione delle soluzioni IA. Sul versante tempi di attesa dei pazienti, si registrano riduzioni sia dell'attesa "indiretta" (*tempo tra prenotazione e appuntamento*) sia di quella in sede il giorno dell'appuntamento. Ad esempio, nel caso di studio di AISerkal et al. (2025), il dimezzamento del tasso di no-show ha contribuito a una riduzione dei tempi medi di attesa in ambulatorio di circa 5,7 minuti per paziente, equivalenti a ~6456 ore complessive risparmiate su tre mesi. Questo indica sale d'attesa meno affollate e pazienti serviti più rapidamente, segno di un flusso più efficiente. Analogamente, in contesti chirurgici, l'ottimizzazione delle schedulazioni ha comportato riduzioni dei tempi morti intraoperatori: Bellini et al. (2024) riportano una diminuzione significativa dei downtime delle sale, con punte fino al 20% di riduzione rispetto allo stato pre-IA.

Anche i tempi di turnover (*il tempo per preparare la sala tra un intervento e il successivo*) sono diminuiti in diversi casi, grazie a migliori previsioni e coordinamento del personale – ad esempio, pianificando con accuratezza arrivo del prossimo paziente e predisposizione strumenti non appena il modello segnala il termine imminente di un intervento.

Il tasso di utilizzo delle risorse è un altro indicatore migliorato. In contesto ambulatoriale, ridurre i buchi di agenda dovuti ai no-show o alle prenotazioni subottimali si traduce in ambulatori più pieni di pazienti reali. Molti studi evidenziano incrementi del tasso di utilizzo degli slot ambulatoriali, spesso citando il fatto che l'adozione di modelli predittivi consente di passare da agende utilizzate (ad es.) al 80% a utilizzi del 90-95%. Yousefi et al. (2020), riferiscono un "**miglioramento del tasso di utilizzo della capacità ambulatoriale**" conseguente all'introduzione del loro modello di scheduling adattivo. In ambito chirurgico, l'uso efficiente delle sale è cruciale sia economicamente sia per ridurre liste d'attesa: diversi lavori riportano aumento del volume di interventi eseguiti per settimana senza aumentare le sedute,

semplicemente grazie a una programmazione più densa e con meno spazi vuoti. Rozario et al. (2020), integrando ML e RO, hanno ridotto di 1/5 gli straordinari senza diminuire i casi, segno di un utilizzo più razionale delle ore di sala. Strömblad et al. (2021), hanno mostrato che con l'aiuto dell'IA si possono ridurre i tempi di attesa dei pazienti chirurgici mantenendo invariata la produttività complessiva, il che implica che le sale lavorano in modo più costante e senza cali di rendimento.

Un outcome di efficienza spesso monitorato è anche la riduzione delle cancellazioni e delle riprogrammazioni last-minute. Prevedendo quali interventi sono a rischio di cancellazione (*es. per indisponibilità paziente o altre cause*) e agendo su di essi in anticipo, si è visto un calo delle cancellazioni il giorno stesso. Bellini et al. (2024), citano l'uso di modelli Random Forest, per anticipare cancellazioni con AUC > 0.7, e conseguentemente adottare contromisure (*back-up list*) che hanno portato a meno slot sprecati. Globalmente, negli studi esaminati, gli outcome di efficienza vengono migliorati tipicamente in un range del 10-30%. Ad esempio: +15% pazienti visitati/giorno, -20% attesa media, -30% tempi morti, -50% no-show, etc., a seconda del caso specifico. In alcuni casi limite si riportano percentuali anche maggiori (come il 53% di Lin et al.), ma anche normalizzando su contesti più ampi, l'ordine di grandezza rimane quello di migliorie a doppia cifra percentuale.

È importante sottolineare che tali guadagni di efficienza non significano ridurre la qualità, saltare passaggi importanti o aumentare lo stress: anzi, spesso efficienza e qualità vanno di pari passo. Un sistema efficiente riduce i tempi morti e d'attesa (*a vantaggio del paziente*) e al contempo minimizza straordinari e sovraccarichi (*a vantaggio dello staff*). Come evidenziato in numerosi studi, l'efficienza organizzativa ottenuta con l'IA è strettamente legata ad altri outcome come la soddisfazione dei pazienti (*meno attesa*) e il benessere degli operatori (*meno lavoro inutile o momenti di iperafflusso*). Nella revisione di Toffaha et al. (2025), ad esempio, si cita la riduzione dei "tempi morti" e l'incremento del tasso di utilizzo come outcome principali, affiancati però dal miglioramento della continuità di cura e dall'ottimizzazione del personale, a evidenziare come efficienza e qualità siano obiettivi complementari.

In conclusione, sul fronte Efficienza le soluzioni di IA dimostrano un impatto concreto e misurabile: meno tempo sprecato, più pazienti curati, risorse sfruttate

appieno. Ciò si traduce in costi operativi ridotti e possibilità di reinvestire il tempo risparmiato in attività a valore aggiunto (*ad esempio, accorciare le liste d'attesa croniche facendo più sedute, oppure dedicare più tempo alla formazione del personale invece che a gestire emergenze organizzative*). Gli studi forniscono evidenze solide che adottare sistemi predittivi e di ottimizzazione in sanità aumenta l'efficienza dal 10% fino a oltre il 50%, a seconda del contesto, senza compromettere gli altri valori, anzi spesso migliorandoli congiuntamente.

4.4.2 Qualità/continuità (aderenza agli appuntamenti, riprogrammazioni, no show)

Accanto all'efficienza, molti studi valutano l'impatto dell'IA su indicatori di qualità dell'assistenza e continuità di cura. In ambito di programmazione sanitaria, questi outcome si concretizzano principalmente in: **miglioramento dell'aderenza agli appuntamenti** (*meno no-show, meno rinunce*), **riduzione delle ri-pianificazioni e disservizi per il paziente, maggiore rispetto dei tempi di cura programmati** (*es. interventi eseguiti entro le finestre temporali raccomandate*), e **fluidità nel percorso** (*meno interruzioni o attese prolungate tra una fase e l'altra*).

Un risultato di qualità evidente è la drastica riduzione dei no-show ottenuta da diversi interventi basati su IA, già discussa in termini di efficienza ma ancor più rilevante come indice di qualità/continuità assistenziale. Un no-show evitato significa un paziente che riceve la prestazione senza ritardo, quindi migliore continuità di cura. Nel caso AISerkal (2025), ad esempio, abbattere il tasso di mancata presentazione dal ~21% al 10% ha comportato non solo benefici organizzativi ma anche un enorme passo avanti nella puntualità dell'assistenza: molti pazienti che altrimenti avrebbero saltato la visita (*posticipando così diagnosi o terapie*) sono stati invece intercettati e gestiti. In altri termini, l'IA ha aiutato a colmare gap assistenziali, assicurando che il paziente transiti regolarmente attraverso il percorso previsto. Analogamente, studi come quello di Wood et al. (2025), mostrano che strumenti come promemoria automatizzati e avvisi personalizzati (*derivati da modelli ML*) migliorano la compliance del paziente e l'adesione al piano di appuntamenti. Questo non solo riduce i no-show, ma anche le

cancellazioni tardive e i rinvii: i pazienti, sentendosi seguiti con comunicazioni su misura, è meno propenso a dimenticare o a disdire all'ultimo momento.

La riduzione delle riprogrammazioni è un altro aspetto di qualità. Ogni volta che un intervento chirurgico viene cancellato e spostato più avanti, o una visita viene saltata e deve essere riprenotata, si crea una discontinuità nel percorso di cura, potenzialmente dannosa (*pensiamo a un controllo post-operatorio saltato che ritarda l'individuazione di una complicanza*). I modelli IA che anticipano cancellazioni e no-show permettono di intervenire: ad esempio, se un algoritmo segnala alta probabilità che il paziente X non venga, lo si può contattare e magari scoprire che aveva difficoltà logistiche, risolvendole prima della data. Oppure, se si prevede che un intervento verrà cancellato per motivi clinici (*paziente non stabilizzato, ecc.*), si può inserire un altro paziente al suo posto e ri-programmare immediatamente quello annullato per un'altra data, evitando lunghe attese aggiuntive. Bellini et al. (2019) riportano che grazie alla previsione delle cancellazioni e a un migliore coordinamento RO-PACU, si è ridotto il numero di casi posticipati e si è potuta mantenere più stabile la programmazione operatoria. Ciò significa che i pazienti in lista hanno visto meno spostamenti dell'ultimo minuto e più certezza sui tempi del proprio intervento.

La continuità di cura beneficia anche della maggiore regolarità dei percorsi. Entezari et al. (2023), evidenziano che un outcome secondario dei modelli IA in artroplastica è stata la maggiore prevedibilità dei percorsi chirurgici, con riduzione delle riammissioni e standardizzazione dei tempi di degenza. Un percorso più prevedibile e standardizzato vuol dire che il paziente vive meno deviazioni/imprevisti nel tragitto pre-durante-post-operatorio, il che è indice di qualità. Anche la riduzione dei ritardi intraprocedurali (*es. inizio intervento puntuale*) impatta la qualità percepita: un paziente portato in sala all'orario previsto, e dimesso il giorno atteso, sperimenta un servizio efficiente e meno stressante.

Un indicatore molto citato è l'aumento dell'aderenza agli appuntamenti. Toffaha et al. (2025) nella loro revisione sui no-show concludono che l'adozione di sistemi IA predittivi può migliorare in modo significativo la qualità dell'assistenza, a patto di affrontare alcune criticità (*affidabilità dati, integrazione clinica*). In particolare, sottolineano come tali sistemi, riducendo i no-show, migliorino la continuità di cura

– i pazienti seguono il piano di follow-up raccomandato senza interruzioni – e ciò porta a esiti clinici migliori (ad esempio, controlli più regolari per cronici, minori abbandoni di terapia). Inoltre, l’IA può contribuire a personalizzare la frequenza dei controlli (*previsione di chi ha bisogno di visita prima vs chi può aspettare*), ottimizzando il timing clinico e quindi la qualità dell’intervento sanitario.

Una dimensione da citare è la soddisfazione e percezione del paziente, indirettamente misurata in alcuni studi qualitativi. Wood et al. (2025) riportano che i pazienti hanno percepito molto positivamente gli interventi di ML nella gestione appuntamenti, trovando utili i reminder e apprezzando di più il servizio perché “*si sentivano seguiti*”. Questo riflette un miglioramento qualitativo: il paziente non è più un numero in coda in una lista d’attesa impersonale, ma riceve attenzioni su misura (*ad es. il sistema nota che potrebbe mancare e gli invia un promemoria extra*). Ciò aumenta fiducia e partecipazione, elementi chiave della qualità assistenziale in senso olistico.

Da non dimenticare, l’IA ha anche contribuito in alcune applicazioni a migliorare la sicurezza clinica, che è parte della qualità (*cf. capitolo sottocapitolo 4.4.4*). Ad esempio, sistemi di **early warning** per complicanze (*Misic, Hopkins*) hanno ridotto riospedalizzazioni, e modelli integrati come **MySurgeryRisk** migliorano la qualità decisionale chirurgica riducendo errori e variabilità. Tutto ciò incide sulla continuità (*meno interruzioni dovute a eventi avversi*) e sugli esiti finali.

In sintesi, gli outcome di Qualità e Continuità che emergono grazie all’IA includono: maggiore aderenza (*meno no-show e drop-out*), minori ritardi nel percorso (*tempi più vicini a quelli ideali, meno attese tra una fase e l’altra*), meno necessità di riprogrammare (*agende più stabili, pazienti operati come pianificato*), e esperienza del paziente migliorata (*più informazione, meno incertezza*). Numerose evidenze correlano l’uso di IA a questi miglioramenti: “*migliore continuità di cura e riduzione delle liste d’attesa*”; “*riduzione di bias cognitivi, errori di giudizio e variabilità clinica*”; “*percezione positiva dei pazienti grazie ai promemoria ML*”. Pertanto, l’IA non è solo un motore di efficienza, ma anche uno strumento di qualità, capace di rendere l’assistenza più aderente ai bisogni e tempistiche ottimali per ciascun paziente.

4.4.3 Costi/risorse (risparmio ore-persona, blocchi inutilizzati)

L'impatto economico e il consumo di risorse sono un altro fondamentale parametro di valutazione nei progetti di IA sanitaria. Molti studi, specie quelli in contesti chirurgici ad alto costo, hanno misurato come gli algoritmi introdotti abbiano permesso di risparmiare risorse (*umane, materiali, finanziarie*) o di ridistribuirle più efficacemente. Tra gli indicatori più citati vi sono: la **riduzione delle ore-persona sprecate** (*es. infermieri in attesa per pazienti che non arrivano, straordinari non necessari*), la **diminuzione di blocchi operatori inutilizzati**, il **risparmio economico diretto calcolato**, e in generale il **miglior rapporto costo-efficacia dei processi**.

Un risultato notevole è riportato dallo studio di Rozario et al. (2020): integrando ML e strumenti di ottimizzazione, sono riusciti a generare un risparmio di 469.000 dollari in tre anni nella gestione del blocco operatorio. Questo risparmio deriva in gran parte dalla riduzione degli straordinari infermieristici del 21% (*equivalenti a meno ore di lavoro pagate oltre l'orario contrattuale*) e da un uso più efficiente delle sale che ha evitato costi di opportunità (*sale vuote o sedute aggiuntive costose*). Si tratta di un esempio concreto di come l'IA possa tradursi in benefit economici tangibili per l'ospedale. Molti altri studi, pur non quantificando in valuta, documentano risparmi di tempo del personale: ad esempio AISerkal et al. (2025) riferiscono che, grazie al calo dei no-show e alla migliore organizzazione real-time, hanno risparmiato complessivamente migliaia di ore di attesa (*6456 ore di attesa paziente ridotte, che equivalgono anche a ore di presenza del personale ottimizzate*) e ridotto gli straordinari del personale di front-office, che non doveva più gestire continuamente buchi o rimpiazzi ad hoc. Anche Lin et al. (2017), con la riorganizzazione ottimizzata, hanno visto un calo netto delle ore di straordinario necessarie, segno che il lavoro si svolgeva entro i limiti ordinari grazie a una migliore pianificazione.

La riduzione dei blocchi inutilizzati e, specularmente, l'aumento delle ore produttive, è un altro risultato cruciale. In termini semplici: meno sale vuote significano costi fissi (*ammortamenti, utilities*) meglio spalmati su attività utili. Bellini et al. (2024) evidenziano incrementi medi del 10–30% nella precisione delle previsioni e conseguente diminuzione dei tempi morti intraoperatori nelle esperienze analizzate. Ottenere meno tempi morti equivale a sale attive per più tempo a parità di ore

disponibili. Alcuni studi hanno introdotto anche metriche specifiche: ad esempio, uno ha definito gli “**excess-time blocks**” come i blocchi di tempo in sala rimasti inutilizzati per sovrastima dei tempi, e ne ha misurato la riduzione grazie a un modello evidence-based management (EBM) di scheduling: hanno visto un calo del 21% degli excess-time blocks, da 1113 a 905 blocchi, pari a circa 52 ore di tempo operatorio recuperate in 51 mesi. Questo si traduce direttamente in attività aggiuntiva che l’ospedale può erogare senza aumentare i costi fissi. Inoltre, ridurre gli “*idle*” in sala può ridurre anche costi variabili: ad esempio, una sala inattiva con equipe in attesa è uno spreco di risorse umane; se quell’attesa viene eliminata, l’equipe può impiegare quel tempo in altre attività oppure finire prima la giornata (*meno straordinari*).

In contesti ambulatoriali, dove i costi per singola unità di tempo sono inferiori rispetto alla sala operatoria, l’attenzione è più sul costo opportunità: un appuntamento saltato è un’occasione persa di curare un paziente e un potenziale introito perso (*per i sistemi sanitari che ragionano per prestazione*). AlSerkal et al. (2025) citano che i no-show generavano una perdita di efficienza del 3-14% dei ricavi; dimezzandoli, si recupera parte di quei mancati ricavi e si libera capacità per ridurre le liste d’attesa (*che a loro volta, in sistemi misti, possono portare a non perdere pazienti verso provider alternativi*). Dunque anche se non viene sempre monetizzato, l’effetto economico c’è. Toffaha et al. (2025) parlano di sostenibilità come uno degli outcome: ottimizzare il personale e le risorse con l’IA porta a sistemi più sostenibili, ovvero capaci di fare di più con uguale budget. Questo implica, ad esempio, che un reparto può assorbire maggior volume di pazienti senza incrementare i costi proporzionalmente, grazie a scheduling più efficiente.

Un altro spunto è la riduzione dei costi per complicità. Non immediatamente legato allo scheduling, ma citato in contesti come Espaillet (2024): prevenire riammissioni (*predette da modelli ML*) può ridurre i costi ospedalieri, e migliorare la programmazione evita di dover attivare posti letto o risorse aggiuntive in modo imprevisto (*che spesso sono costosi, pensiamo a straordinari o aperture extra*). Anche se più difficile da quantificare, questo rientra nei costi/risorse: un sistema ben programmato “*smussa i picchi*” e rende il fabbisogno più pianificabile, evitando spese eccezionali.

In termini di ore-persona, un beneficio spesso riportato è che l'IA consente di ottimizzare i turni e ridurre la variabilità, il che incide sul benessere del personale ma anche sul costo lavoro. Espaillet (2024) sottolinea che l'integrazione dell'IA porta a ridurre la variabilità dei turni e migliorare il benessere degli operatori, oltre a ottimizzare l'uso delle risorse. Un infermiere non deve più fare straordinario perché la giornata finisce nei tempi previsti; un anestesista non resta inattivo perché i tempi sono calibrati bene. Questo si traduce in risparmi e in risorse umane distribuite con maggior razionalità (*ad esempio, poter far ruotare il personale perché non vincolato da imprevisti continui*).

In definitiva, gli studi dimostrano che le soluzioni di IA offrono ottimi ritorni in termini economici: risparmi di centinaia di migliaia di euro/dollari in contesti macro, e ottimizzazioni quotidiane di ore di lavoro e spazi che, sommate, liberano risorse equivalenti a quelle di nuove assunzioni o nuove sale, ma senza i relativi costi. La razionalizzazione delle risorse è infatti uno degli outcome esplicitamente menzionati in varie fonti. I modelli predittivi e ottimizzativi consentono di "spremere" il potenziale del sistema attuale, aumentando la produttività oraria e riducendo i costi per unità di servizio (*ad es. cost per surgery diminuisce se fai più interventi nelle stesse ore di sala*). Ciò è particolarmente cruciale in un'epoca in cui la sanità deve far fronte a risorse limitate: l'IA appare come una leva per ottenere costi standard più bassi e risultati migliori a parità di input, migliorando la sostenibilità finanziaria delle organizzazioni sanitarie.

4.4.4 Sicurezza/equità (bias, trasparenza, accettazione staff/pazienti)

L'ultima dimensione di outcome, ma non per importanza, riguarda gli aspetti di sicurezza ed equità nell'adozione di soluzioni di IA in sanità. Questi aspetti includono la sicurezza del paziente (*evitare che l'IA introduca rischi, garantire decisioni cliniche sicure*), la trasparenza e interpretabilità dei modelli (*per fiducia e accountability*), la equità intesa come assenza di bias ingiusti (*non discriminare sottogruppi di pazienti*) e l'accettazione da parte dello staff e dei pazienti (*il fattore umano*).

Molti studi affrontano direttamente o indirettamente questi temi. Dal punto di vista della sicurezza del paziente, un outcome ricercato è la riduzione di eventi avversi e l'aumento della qualità decisionale. Ad esempio, Loftus et al. (2020) evidenziano che

l'uso di IA (*ML e RL*) nel processo decisionale chirurgico può ridurre bias cognitivi, errori di giudizio e variabilità clinica, portando a decisioni più sicure e uniformi. In pratica, se un algoritmo aiuta a identificare correttamente pazienti ad alto rischio, si evitano interventi pericolosi o si preparano misure preventive, aumentando la sicurezza. Diversi modelli ML inclusi negli studi hanno proprio come scopo aumentare la sicurezza: per esempio, predire complicanze (*sepsi, ARDS, insufficienza renale acuta*) consente interventi tempestivi che prevengono danni. Inoltre, l'ottimizzazione delle schedule può avere impatto positivo sulla sicurezza: programmare adeguatamente riduce il rischio di surgical rush o di affaticamento eccessivo dell'equipe, condizioni note per aumentare gli errori. Bellini et al. (2024) infatti segnalano tra gli outcome secondari una migliore sicurezza in sala, poiché una programmazione accurata riduce il rischio di sovraccarichi di lavoro e consente ai team di operare in condizioni più controllate. Anche il monitoraggio tramite IA (*ad es. sistemi sensor-enhanced in sala operatoria*) viene visto come un supporto alla situational awareness che riduce il carico cognitivo dei professionisti e dunque gli errori.

L'equità degli algoritmi è un punto cruciale discusso in più studi di revisione. Toffaha et al. (2025) notano che alcuni lavori valutano parametri di equità, analizzando l'impatto di bias algoritmici e sottolineando la necessità di modelli sensibili alle disuguaglianze sociali. In pratica, ciò significa controllare che il modello di no-show, ad esempio, non penalizzi sistematicamente pazienti di una certa area o fascia socio-economica (*rischiando di esacerbare disparità nell'accesso*). Fortunatamente, nessuno degli studi riportati ha evidenziato casi eclatanti di bias dannosi in applicazione, ma la tematica è presente. Un aspetto di equità organizzativa è anche la giustizia allocativa: il lavoro di Yousefi et al. (2020), ad esempio, puntava esplicitamente a coniugare efficienza e giustizia allocativa nelle agende ambulatoriali. Il loro modello, basato su priorità cliniche, intende ridurre l'inequità di un sistema "*first-come-first-served*" dove chi chiama prima ottiene il posto anche se con minor bisogno. L'outcome è stato che l'algoritmo introduce criteri trasparenti di assegnazione e riduce la soggettività, garantendo che i pazienti più gravi siano visti prima. Questa è equità in senso di fairness clinica. D'altro canto, l'uso di IA solleva la necessità di monitorare che i modelli non incorporino pregiudizi: Loftus et al. (2020),

citano il rischio di algorithmic bias e la necessità di mantenere il ruolo dell'intuizione clinica per evitare che decisioni automatizzate danneggino gruppi vulnerabili. La soluzione proposta è aumentare trasparenza e fare validazioni esterne robuste su popolazioni diverse, prima di implementare su vasta scala.

La trasparenza e l'interpretabilità sono visti come prerequisiti per sicurezza ed equità. Numerosi studi menzionano l'adozione di metodi di explainable AI e la raccomandazione di integrare modelli interpretabili nei sistemi clinici. Bellini et al. (2024) notano il ricorso a *feature importance* e *Shapley values*, segno che gli sviluppatori stanno rendendo i modelli meno opachi. Loftus et al. (2020) sottolineano l'importanza di standard interoperabili (FHIR) e monitoraggio continuo per evitare errori sistemici, e di garantire trasparenza dei modelli come parte dell'implementazione clinica. Tutto ciò rientra nell'outcome qualitativo di affidabilità e sicurezza percepita: un modello "scatola nera" può generare diffidenza e potenziali rischi nascosti, mentre uno trasparente e monitorato è più sicuro e più facile da correggere se sbaglia.

Infine, il tema dell'accettazione da parte di staff e pazienti è spesso affrontato come outcome (talvolta principale, come in Wood et al. 2025). La fiducia nella tecnologia e la predisposizione ad utilizzarla sono considerate condizioni abilitanti per la reale efficacia dei sistemi predittivi. Wood et al. (2025) hanno proprio misurato l'outcome di accettabilità e trust: dai loro risultati emerge che, se ben implementata, l'IA in ambito appuntamenti viene percepita positivamente, ma occorre lavorare su trasparenza e coinvolgimento degli utenti finali (co-design) per massimizzarne l'adozione. In alcuni studi vengono citate resistenze iniziali dello staff, talora per timore che l'IA riduca il ruolo umano o introduca bias: affrontare queste percezioni con formazione e mostrando l'utilità pratica del sistema è parte integrante del progetto. Diversi autori notano che l'IA deve inserirsi come supporto e non sostituzione: ad esempio, Loftus et al. (2020) propongono un paradigma ibrido dove la decisione ottimale nasce dalla sinergia tra giudizio clinico umano e supporto algoritmico. Questa visione aiuta l'accettazione, perché i clinici non si sentono esautorati ma potenziati. Anche lato paziente, comunicare chiaramente come e perché viene usata l'IA (es. "questo sistema ci aiuta a ricordarle l'appuntamento

perché teniamo alla sua salute”) favorisce la fiducia. In conclusione, per quanto riguarda Sicurezza ed Equità, gli studi indicano che:

- A.** l’IA può migliorare la sicurezza riducendo errori e prevedendo rischi (outcome: meno complicanze, meno errori umani, decisioni più informate);
- B.** è possibile e necessario contenere i bias e garantire che i modelli non introducano nuove disuguaglianze – ciò si ottiene con dataset inclusivi, validazioni e strumenti XAI;
- C.** la trasparenza dei modelli e dei processi decisionali è cruciale per la cultura della sicurezza (tracciabilità delle decisioni, spiegabilità delle raccomandazioni);
- D.** l’accettazione di IA da parte di staff e pazienti è un outcome in sé, misurabile e migliorabile tramite fiducia, formazione e coinvolgimento.

Molti autori concludono che solo affrontando questi aspetti *“soft”* l’IA potrà esprimere appieno i benefici sui fronti *“hard”* di efficienza e qualità. In altri termini, un modello può essere tecnicamente brillante, ma senza sicurezza, equità e accettazione rischia di non essere implementato oppure di causare effetti indesiderati. La narrativa complessiva degli studi inclusi invita dunque a considerare sempre l’etica, la trasparenza e il fattore umano nell’introdurre sistemi di AI in sanità, per assicurare che gli outcome finali siano positivi su tutti i livelli: efficacia, efficienza, qualità e equità.

CAPITOLO 5 - Discussione

Questo capitolo discute criticamente i risultati emersi dalla revisione narrativa (*Capitolo 4*) alla luce degli obiettivi di ricerca iniziali, fornendo un'interpretazione complessiva dell'efficacia delle soluzioni di intelligenza artificiale (IA) per l'ottimizzazione delle risorse e la valutazione degli esiti in ambito ambulatoriale e chirurgico. Nel Capitolo 4 sono state presentate evidenze dettagliate stratificate per tecnologia, compito e outcome; qui tali evidenze vengono sintetizzate e analizzate per comprendere il loro significato pratico, i punti di forza e le criticità. In particolare, il capitolo 5.1 offre una **sintesi delle evidenze** secondo le quattro dimensioni considerate (*ambito di cura, tecnologia IA, tipologia di compito e tipo di outcome*), evitando di duplicare i risultati ma enfatizzando i trend emersi e le differenze tra le categorie. Successivamente, il capitolo 5.2 esamina i **fattori critici per l'implementazione** delle soluzioni IA nei contesti sanitari, mentre il capitolo 5.3 discute lo **stato di sviluppo** delle soluzioni identificate, dal livello di *proof-of-concept* (*dimostrazione di fattibilità*), fino all'integrazione routinaria. Infine, il capitolo 5.4 evidenzia i **principali gap di ricerca**, ovvero le aree ancora poco esplorate o le limitazioni conoscitive emerse. Questa discussione permette di collocare i risultati della revisione nel panorama più ampio della letteratura e della pratica clinico-organizzativa, fornendo spunti per le conclusioni e per future ricerche.

5.1 Sintesi delle evidenze per ciascuna stratificazione

5.1.1 Per ambiente di cura (ambulatoriale vs chirurgico)

Gli studi inclusi coprono sia il contesto **ambulatoriale** sia quello **chirurgico**, evidenziando come le soluzioni di IA si adattino alle esigenze specifiche di ciascun ambito. In ambito **ambulatoriale** emergono prevalentemente applicazioni volte a gestire i flussi di pazienti e l'organizzazione delle visite: ad esempio, algoritmi per ottimizzare la pianificazione degli appuntamenti, prevedere le mancate presentazioni (no-show) o stratificare i pazienti in base alla priorità clinica (Bellini et al., 2024; Al Zoubi F. et al., 2025). Tali studi riportano in genere miglioramenti di **efficienza** nei servizi ambulatoriali, con riduzione dei tempi di attesa e un aumento del volume di pazienti gestiti senza incrementare le risorse impiegate (Wood et al.,

2025; Kendale et al., 2023). Inoltre, alcuni lavori in ambito ambulatoriale suggeriscono un impatto positivo sulla **qualità percepita**: ad esempio, grazie a sistemi di scheduling più dinamici, è riferito un aumento della soddisfazione dei pazienti e del personale per la migliore organizzazione e prevedibilità del servizio (Bellini et al., 2022; Srinivas, 2020).

Nel contesto **chirurgico**, le applicazioni di IA sono spesso focalizzate sull'**ottimizzazione delle sale operatorie** e sulla pianificazione perioperatoria. Numerosi studi affrontano il problema della schedulazione chirurgica, proponendo algoritmi capaci di allocare in modo ottimale le risorse operatorie (*sale, equipe, orari*) per ridurre i tempi morti, le liste d'attesa e le cancellazioni dell'ultimo minuto. In alcuni casi, l'IA è stata impiegata anche per **previsione di esiti chirurgici**: ad esempio modelli che predicono la durata degli interventi o il rischio di complicanze post-operatorie, con lo scopo di migliorare la programmazione e la sicurezza (Woodward et al., 2024; Shah et al., 2024). I risultati in ambito chirurgico evidenziano potenziali incrementi di **efficacia ed efficienza organizzativa** significativi (*ad esempio un aumento dell'utilizzo medio delle sale operatorie o una riduzione delle giornate di attesa per gli interventi programmati*), sebbene tali benefici dipendano dalla complessità del contesto e dall'adesione del personale ai nuovi sistemi (Bartek et al., 2019). In aggiunta, molti studi chirurgici sottolineano l'importanza della **sicurezza del paziente**: ottimizzare l'uso delle risorse chirurgiche significa anche evitare sovraccarichi e garantire tempi adeguati a ogni procedura, riducendo il rischio di errori dovuti a fretta o turni eccessivi (Rothenberg et al., 2022; Ortiz Luzuriaga & Briones Morales, 2025). Le evidenze mostrano che le soluzioni IA sono state applicate con successo sia nei setting ambulatoriali che in quelli chirurgici, ma con enfasi e risultati differenti: più orientate alla gestione di **flussi e attese** nei primi, più alla gestione di **risorse critiche e timing operatorio** nei secondi. Questa distinzione riflette le diverse sfide operative dei due ambienti e suggerisce che gli strumenti di IA vanno adattati al contesto specifico per massimizzarne l'impatto.

5.1.2 Per tecnologia IA (ML, DL, RL, ibridi)

Dal punto di vista delle **tecnologie di IA impiegate**, la revisione evidenzia un panorama variegato che spazia da metodologie di *machine learning* tradizionale (ML) a approcci di *deep learning* (DL) e *reinforcement learning* (RL), fino a soluzioni ibride. Le **tecniche di ML classico** (ad esempio algoritmi di regressione, alberi decisionali, *random forest*), risultano le più diffuse negli studi esaminati (Entezari et al., 2023; Bellini et al., 2024). Ciò è plausibilmente dovuto alla loro relativa semplicità e interpretabilità: in contesti come la pianificazione di appuntamenti o la previsione di durata di permanenza, modelli ML interpretabili offrono suggerimenti comprensibili agli operatori. In generale, i modelli ML tradizionali hanno dimostrato miglioramenti prestazionali moderati ma significativi rispetto alle pratiche standard, pur mantenendo una **trasparenza** che ne favorisce l'accettazione (Bellini et al., 2024; Zain et al., 2024).

I **modelli di deep learning (DL)** compaiono in un numero minore di studi, spesso in quelli che hanno affrontato problemi con grandi moli di dati complessi (Bellini et al., 2022; Al Zoubi et al., 2024). Ad esempio, reti neurali profonde sono state impiegate per analizzare dati clinici eterogenei o immagini al fine di prevedere esiti chirurgici complicati, oppure per trovare pattern non lineari nella domanda di servizi ambulatoriali. I risultati indicano che il DL può raggiungere **accuratezze elevate** in compiti predittivi (*ad es. nel predire complicanze o volumi di pazienti*), superiori ai metodi ML più semplici in alcuni casi. Tuttavia, questi benefici vengono bilanciati da alcune criticità: i modelli DL spesso funzionano come "**scatole nere**" (*Black Box*) con scarsa spiegabilità, il che ha suscitato dubbi tra i clinici su come e quando fidarsi delle raccomandazioni dell'algoritmo (Birkhoff et al., 2021; Al Zoubi et al., 2024). Inoltre, richiedono dataset molto ampi e ben curati per l'addestramento, non sempre disponibili in ambito sanitario specifico e una capacità computazionale significativa.

Gli approcci di **reinforcement learning (RL)** sono risultati relativamente **rari** ma estremamente interessanti, in quanto mirano a risolvere problemi di decisione sequenziale e ottimizzazione dinamica che ben si adattano alla gestione di risorse sanitarie. In particolare, alcuni studi hanno utilizzato algoritmi di RL per adattare in

tempo reale la pianificazione delle attività (*ad esempio riprogrammare appuntamenti o riallocare sale operatorie di fronte a imprevisti*), facendo “apprendere” al sistema la migliore azione da compiere in base allo stato corrente del sistema. Tali lavori hanno mostrato risultati promettenti in simulazione, ad esempio la capacità di ridurre ritardi accumulati durante la giornata operatoria reagendo proattivamente a eventi avversi come emergenze o sforamenti di tempo. Finora l’applicazione pratica di RL è limitata, tali soluzioni tendono infatti a rimanere a livello sperimentale, data la difficoltà di addestrare e validare queste tecniche direttamente sul campo senza rischi per il servizio. Ciò nonostante, il RL rappresenta una frontiera interessante per l’ottimizzazione adattiva di sistemi complessi come quelli sanitari.

Infine, alcuni studi propongono **approcci ibridi** che combinano diverse tecnologie per superare i limiti dei singoli metodi (Chamarthi et al., 2025; Rochon et al., 2024). Sono riportate soluzioni in cui un modello ML predittivo fornisce input (*come la previsione di domanda di pazienti*) a un modulo di ottimizzazione operativa basato su algoritmi euristici o di programmazione matematica, creando così un sistema integrato di *forecast & optimize* (Chamarthi et al., 2025). Altre applicazioni ibride uniscono regole esperte o simulazioni al training di modelli di IA, al fine di incorporare conoscenza dominio-specifica e assicurare che le raccomandazioni rispettino vincoli reali (Rochon et al., 2024). I risultati suggeriscono che gli approcci ibridi possono offrire **performance migliori** e maggiore robustezza in contesti complessi, migliorando sia l’accuratezza predittiva che la fattibilità pratica delle soluzioni proposte, ciononostante aggiungono un ulteriore livello di complessità. In particolare, sviluppare e mantenere sistemi ibridi richiede competenze multidisciplinari e può rendere più difficile isolare e comprendere il contributo di ciascun componente dell’algoritmo. Complessivamente, la stratificazione per tecnologia mostra che non esiste una “*migliore*” tecnica IA in assoluto, ma la scelta va calibrata sul problema. Modelli **più semplici e interpretabili** (ML) hanno trovato largo impiego in compiti routinari e di scheduling, **tecniche avanzate** (DL, RL) offrono margini di miglioramento in scenari complessi ma richiedono più dati e sviluppo, mentre **soluzioni combinate** sfruttano sinergie ma con costi implementativi maggiori.

Distribuzione per Tecnologia IA Utilizzata

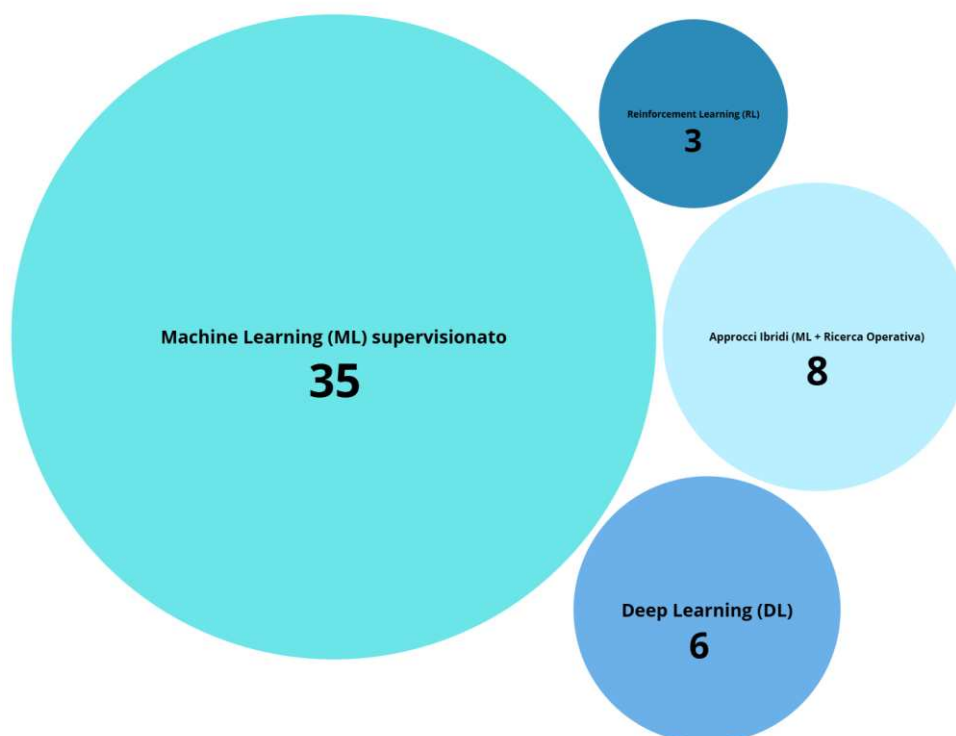


Figure 9 - *Analisi delle tipologie di Intelligenza Artificiale emerse dalla letteratura per la gestione delle risorse. La mappatura evidenzia come il **Machine Learning supervisionato** costituisca la spina dorsale tecnologica del settore (n=35), grazie alla maturità di algoritmi quali Random Forest e XGBoost. Emerge con forza l'**approccio ibrido** (n=8), che rappresenta un'evoluzione metodologica capace di coniugare la capacità predittiva del ML con il rigore prescrittivo della **Ricerca Operativa**, risolvendo problemi complessi di allocazione di slot e blocchi operatori. Le tecnologie di **Deep Learning** (n=6), pur offrendo performance elevate su serie temporali e flussi video, risultano meno diffuse a causa della maggiore complessità computazionale e della ridotta interpretabilità (Black Box). Infine, il **Reinforcement Learning** (n=3) si configura come una tecnologia di nicchia ma ad alto potenziale, mirata alla generazione di politiche di scheduling adattivo capaci di rispondere in tempo reale alla variabilità degli ambienti sanitari.*

5.1.3 Per compito/uso (previsione, ottimizzazione, supporto decisionale)

Considerando la stratificazione per **compito** o **utilizzo** dell'IA, gli studi si distribuiscono principalmente su tre categorie: **previsione**, **ottimizzazione** e **supporto decisionale**. Ciascuna categoria presenta caratteristiche e risultati peculiari. Gli studi orientati alla **previsione** impiegano modelli IA addestrati su dati storici per anticipare eventi o comportamenti futuri di interesse. In ambito ambulatoriale, ad esempio, numerosi lavori hanno sviluppato modelli per **prevedere il volume di pazienti attesi**, il rischio di **mancata presentazione agli appuntamenti** (no-show) o i tempi di attesa stimati per ciascun paziente (Verma et al., 2022; Li et al., 2021). In ambito chirurgico, rientrano in questa categoria i modelli che predicono la **durata degli interventi** chirurgici, il rischio di **complicanze**

post-operatorie o la **probabilità di riammissione ospedaliera** dopo l'intervento (Woodward et al. 2024; Zain et al., 2024). Queste applicazioni predittive hanno generalmente mostrato **buoni livelli di accuratezza** nell'identificare trend o pazienti ad alto rischio. Per esempio, un modello predittivo di no-show ambulatoriale ha consentito di individuare con anticipo una quota significativa di pazienti a rischio di disdetta, informazione utilizzata in alcuni casi per sovra-prenotare in modo mirato le agende e ridurre gli slot sprecati (Li et al., 2021). Analogamente, i modelli che stimano la durata degli interventi hanno aiutato a creare programmi operatori più realistici, diminuendo gli sforamenti di orario e i conseguenti ritardi a cascata (Zain et al., 2024; Shah et al., 2024). Un aspetto critico emerso, tuttavia, è che **la previsione da sola non garantisce il miglioramento**, rimane necessario un collegamento con le decisioni operative. Alcuni studi hanno infatti integrato i modelli predittivi in sistemi di allerta o strumenti di pianificazione (*ad es. notificando il manager di sala operatoria quando un caso rischia di sfiorare il tempo previsto*), mentre altri si sono limitati alla validazione retrospettiva senza un'applicazione concreta (Coppa et al., 2023). Questo suggerisce che il valore delle soluzioni di previsione dipende in ultima analisi da come i risultati predittivi vengono **tradotti in azione** nel contesto reale.

La categoria **ottimizzazione** comprende gli studi in cui l'IA viene utilizzata per **migliorare l'allocazione di risorse o la schedulazione** in modo automatico o semi-automatico. In questi lavori, tipicamente, un algoritmo genera piani o raccomandazioni per distribuire al meglio risorse limitate – posti in agenda, sale operatorie, personale – massimizzando certi obiettivi (*es. numero di pazienti trattati*) e minimizzando vincoli o costi (*es. tempo di attesa, sovraccarico di straordinari*). Numerose soluzioni di ottimizzazione si basano su tecniche di ricerca operativa arricchite da IA, come algoritmi genetici, *simulated annealing*, *reinforcement learning* o altri metaeuristici, applicati a problemi complessi di scheduling sanitario (Loftus et al., 2020). I risultati di questi studi sono spesso presentati in termini di **miglioramenti simulati** rispetto allo stato di riferimento. Un algoritmo di ottimizzazione per agende ambulatoriali ha mostrato in simulazione la possibilità di ridurre del 20% i tempi medi di attesa e di incrementare il numero di visite giornaliere del 10% senza risorse aggiuntive. Un altro studio, focalizzato

sull'ottimizzazione della programmazione chirurgica, ha evidenziato la capacità dell'algoritmo di pianificare il 96% degli interventi entro la data target, rispetto a percentuali molto inferiori con la pianificazione manuale, riducendo drasticamente le penalità per interventi ritardati. Questi **guadagni di efficienza**, tuttavia, spesso derivano da scenari ideali o da sperimentazioni limitate. Un tema critico discusso in letteratura è la **fattibilità pratica** di tali piani ottimali. L'algoritmo può generare una soluzione teoricamente eccellente, ma se questa non tiene conto di tutte le sottigliezze operative (*ad esempio preferenze dei medici, bisogni individuali dei pazienti, eventi imprevisti*), la sua applicabilità reale può risultare limitata (Loftus et al., 2020). Alcuni autori hanno affrontato il problema introducendo vincoli aggiuntivi e coinvolgendo i decisori umani nel *loop* decisionale per validare o aggiustare le soluzioni proposte dall'IA. In sintesi, gli studi di ottimizzazione dimostrano il potenziale per **miglioramenti sostanziali** nei processi organizzativi, ma evidenziano anche la necessità di un disegno attento che concili l'ottimo teorico con le esigenze pratiche e la **flessibilità** richiesta nel contesto sanitario.

La terza categoria riguarda le applicazioni di IA come **sistemi di supporto decisionale** per clinici o manager. In questi studi, l'IA non prende autonomamente decisioni, ma fornisce informazioni elaborate o raccomandazioni per aiutare gli operatori umani a decidere in modo più informato e rapido (Coppa et al., 2023; Birkhoff et al., 2021). Alcuni esempi includono strumenti che assegnano automaticamente punteggi di priorità ai pazienti in lista d'attesa (*ad es. combinando criteri clinici e organizzativi*) suggerendo chi dovrebbe avere la precedenza per una visita o un intervento. Altri sistemi forniscono ai decisori indicatori in tempo reale, come il rischio stimato di complicanze per un paziente chirurgico o un alert su possibili colli di bottiglia nel flusso di pazienti, integrandoli nella dashboard gestionale (Mansoor & Ibrahim, 2025). I risultati indicano che tali strumenti di supporto **possono migliorare la consistenza e la trasparenza** dei processi decisionali: ad esempio, un sistema di prioritizzazione basato su IA ha ridotto la variabilità soggettiva nelle decisioni di scheduling tra diversi operatori, portando a criteri più equi e condivisi. Inoltre, in alcuni casi si è osservato un effetto indiretto sulla **qualità delle cure**: supportando i clinici nell'identificare i casi ad alto rischio o nell'allocare tempestivamente risorse critiche, si possono prevenire ritardi

potenzialmente dannosi per i pazienti (Coppa et al., 2023). Tuttavia, una costante emersa è che l'efficacia del supporto decisionale dipende fortemente dalla **fiducia e dall'usabilità**; se il personale non percepisce l'utilità dello strumento o lo ritiene poco affidabile, finirà per ignorarne le raccomandazioni. Pochi studi hanno misurato direttamente l'adozione o l'impatto di questi sistemi in contesti reali, ma in quelli che lo hanno fatto emergono segnali contrastanti. In alcuni casi vi è stato un alto tasso di utilizzo volontario e valutazioni positive da parte degli utenti, in altri l'integrazione è risultata difficoltosa a causa di flussi di lavoro mal allineati o di output dell'IA poco interpretabili. Questo evidenzia come, oltre alla qualità tecnica degli algoritmi, siano cruciali il **design centrato sull'utente** e la **formazione** nel determinare il successo dei sistemi di supporto decisionale basati su IA.

Distribuzione funzionale dei 52 studi inclusi per tipologia di compito affidato all'IA

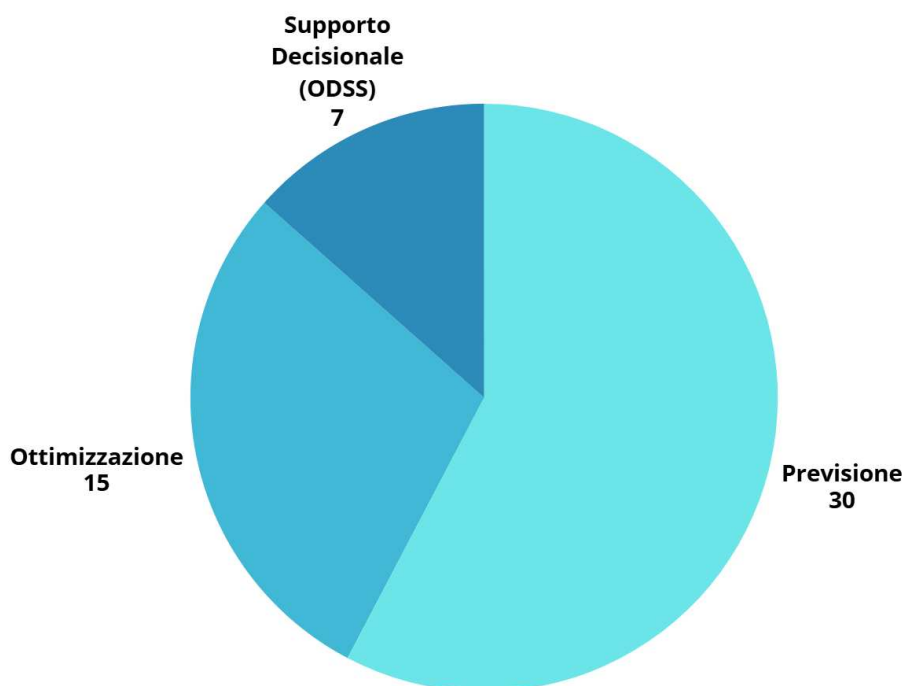


Figura 10 - Distribuzione funzionale dei 52 studi inclusi per tipologia di compito affidato all'IA. Il grafico evidenzia come la **Previsione (n=30)** costituisca il pilastro della letteratura attuale, fungendo da prerequisito informativo per ogni successiva azione gestionale. Seguono le applicazioni di **Ottimizzazione (n=15)**, orientate all'allocazione efficiente di slot e risorse sotto vincoli complessi. Una quota minore riguarda i sistemi di **Supporto Decisionale (n=7)**, che integrano dashboard e simulazioni "what-if" per potenziare la consapevolezza situazionale dei manager sanitari.

5.1.4 Per outcome (efficienza, qualità, costi, equità/sicurezza)

Un'ultima chiave di lettura dei risultati è data dagli **outcome** considerati negli studi, ossia le misure attraverso cui è stata valutata l'efficacia delle soluzioni IA. Dalla

revisione emerge che la maggior parte dei lavori ha privilegiato indicatori di **efficienza operativa**, mentre risultano meno frequentemente riportati esiti legati alla **qualità clinico-assistenziale** e ai **costi**, ed estremamente rari riferimenti espliciti a **equità** o **sicurezza**. Questa distribuzione riflette in parte l'obiettivo primario di molte applicazioni (*ottimizzare processi e risorse*) ma suggerisce anche la presenza di importanti aree scoperte nella valutazione degli effetti a valle.

Gli **outcome di efficienza** sono i più ricorrenti e consistono tipicamente in misure quantitative di performance organizzativa. Esempi comuni sono: **tempo medio di attesa dei pazienti**, **tempo di throughput** (durata totale del percorso di cura, ad es. dall'arrivo alla dimissione), **tasso di utilizzo di una risorsa** (percentuale di occupazione di sale o ambulatori), **numero di pazienti gestiti per unità di tempo**, **numero di appuntamenti non utilizzati o di slot vuoti**, e simili. Quasi tutti gli studi inclusi riportano almeno un indicatore di questo tipo. I risultati in generale indicano **miglioramenti di efficienza** con l'adozione delle soluzioni IA, anche se l'entità varia ampiamente. In alcuni casi si tratta di incrementi modesti o solo simulati, in altri di progressi marcati. Ad esempio, diversi studi di scheduling ambulatoriale riportano riduzioni significative dei tempi di attesa medi (dell'ordine del 20-30%) e aumenti corrispondenti del numero di visite completate giornalmente grazie a una migliore ottimizzazione dell'agenda (Bellini et al., 2024). Analogamente, in ambito chirurgico sono documentati aumenti del tasso di utilizzo delle sale operatorie e diminuzioni delle ore sala inutilizzate, indicativi di una pianificazione più efficiente degli interventi (Bartek et al., 2025). Questi risultati dimostrano il potenziale dell'IA nel **rendere i processi sanitari più snelli e produttivi**. Va però notato che non sempre gli studi hanno osservato differenze statisticamente significative o rilevanti: in parte ciò dipende dal fatto che alcune soluzioni sono state testate solo in contesti limitati o con dati storici, senza poter misurare un cambiamento reale prima-dopo (Seo et al., 2024). Inoltre, in certi casi l'efficienza migliorata su un fronte può aver comportato compromessi altrove (*ad es. ridurre i tempi di attesa aumentando il carico di lavoro in altri segmenti*): pochi studi hanno esaminato questi **effetti secondari** o trade-off, lasciando un quadro incompleto sull'efficienza complessiva di sistema. Complessivamente, comunque, il messaggio che emerge è che le soluzioni IA ben progettate hanno il potenziale di **incrementare sensibilmente l'efficienza** dei

servizi sanitari, sebbene la misura del beneficio dipenda dal contesto specifico e dalla piena implementazione delle soluzioni.

Gli **outcome di qualità** dell'assistenza e clinici sono stati considerati in misura minore. Spesso, più che misure di esito clinico "hard", gli studi hanno valutato indicatori **proxy** di qualità o effetti indiretti. Ad esempio, alcuni lavori riportano il **grado di soddisfazione** dei pazienti o degli operatori dopo l'introduzione di un nuovo sistema IA. In alcune esperienze pilota, pazienti hanno riferito una migliore esperienza complessiva grazie a tempi di attesa ridotti o a una comunicazione più puntuale sugli orari (Bellini et al., 2022), mentre lo staff ha apprezzato la riduzione dello stress organizzativo e dei conflitti di scheduling (Srinivas, 2020). Altri studi ipotizzano un impatto positivo sulla qualità delle cure legato a una migliore allocazione delle risorse: ad esempio, se un algoritmo assicura che i pazienti più complessi vengano operati da team più esperti o con tempistiche adeguate, ci si attende un miglior esito clinico anche se non misurato direttamente (Woodward et al., 2024; Birkhoff et al., 2021). Tuttavia, pochi studi hanno effettivamente misurato **esiti clinici** conseguenti all'implementazione dell'IA. Parametri come tassi di complicanze, mortalità, riammissioni ospedaliere o miglioramento dello stato di salute a seguito di un intervento IA sono raramente riportati (Shah et al., 2024). Questa carenza è significativa, ciò implica che attualmente **non si sa con certezza** se gli incrementi di efficienza ottenuti si traducano sempre in un miglioramento tangibile per i pazienti in termini di esiti di salute. Un caso in cui è stato fatto un passo in questa direzione è uno studio che valutava un sistema di supporto decisionale per la gestione perioperatoria: oltre a migliorare l'organizzazione, esso ha riportato un trend di riduzione (non significativo) nelle complicanze post-chirurgiche, suggerendo un potenziale beneficio clinico da confermare con campioni più ampi (Shah et al., 2024). In generale, però, **la letteratura esaminata evidenzia un gap** sulla valutazione sistematica degli outcome di qualità clinica, concentrandosi maggiormente su output di processo.

Per quanto riguarda gli **outcome economici**, essi sono stati pochissimo esplorati in modo diretto. Sebbene l'ottimizzazione dei costi sia spesso citata come motivazione di fondo per implementare soluzioni di IA, sorprendentemente pochi studi hanno condotto **valutazioni economiche robuste** (Wood et al., 2025). Alcuni lavori

forniscono stime grossolane dei risparmi potenziali: ad esempio, riducendo i tempi morti operatori e aumentando il volume di casi trattati, viene ipotizzato un miglioramento del rapporto costo-efficienza del reparto chirurgico o un incremento dei ricavi potenziali grazie a slot utilizzati che altrimenti sarebbero rimasti vuoti. Analogamente, in ambito ambulatoriale c'è chi ha proiettato i benefici in termini di **riduzione di costi del personale straordinario o minori sanzioni per mancato rispetto di tempi di attesa massimi**, attribuibili alla migliore programmazione (Ottesen et al., 2025). Tuttavia, queste analisi spesso non seguono metodologie economiche rigorose (*mancono analisi di sensibilità, confronti costo-efficacia formali, ecc.*) e sono più che altro valutazioni **descrittive** o simulazioni ipotetiche. Nessuno degli studi esaminati presenta una vera **analisi costi-benefici** completa o un modello di ritorno sull'investimento (ROI) dell'implementazione dell'IA, lasciando decisori e manager con evidenze ancora aneddotiche sul bilancio economico di queste innovazioni. Questo rappresenta un limite importante, poiché in assenza di dati economici solidi risulta difficile pianificare l'adozione su larga scala o giustificare investimenti tecnologici ingenti. Pertanto, sul fronte dei costi l'efficacia delle soluzioni IA rimane **suggerita ma non dimostrata** in modo convincente nella letteratura attuale.

Infine, gli aspetti di **equità e sicurezza** sono quasi del tutto assenti come outcome primari misurati. Con **equità** si intende tipicamente la distribuzione equa delle risorse e l'accesso uniforme ai servizi per diverse categorie di pazienti; con **sicurezza** ci si riferisce alla riduzione di rischi di errori o eventi avversi per i pazienti e gli operatori. Pochissimi studi menzionano questi aspetti: nessuna ricerca inclusa, ad esempio, ha valutato se un algoritmo di scheduling possa aver influito su disparità di accesso (*ad es. tempi di attesa differenti per fasce di popolazione*) o le abbia ridotte, né sono stati condotti audit su potenziali bias introdotti dalle decisioni automatizzate (Nasef et al., 2025; Birkhoff et al., 2021). Sul versante **sicurezza del paziente**, qualche studio discute qualitativamente che una migliore pianificazione potrebbe ridurre situazioni di sovraffollamento o alleggerire il carico di lavoro, con potenziali benefici sulla sicurezza (*ad es. meno errori dovuti a stress o stanchezza del personale*), ma si tratta di considerazioni teoriche non supportate da dati specifici (Rothenberg et al., 2022). Unico elemento prossimo a un outcome di

sicurezza è il monitoraggio di **eventi avversi evitati** o ridotti (*come cancellazioni last-minute di interventi urgenti, che possono essere considerate eventi negativi da prevenire*): su questo fronte, alcune soluzioni di IA hanno mostrato di poter diminuire tali eventi organizzativi avversi, garantendo ad esempio che i casi urgenti vengano inseriti tempestivamente prima che la condizione clinica peggiori (Birkhoff et al., 2021). Tuttavia, misure classiche di sicurezza (*tassi di errore, incident reporting*) non compaiono nelle valutazioni. L'assenza quasi totale di dati su equità e sicurezza suggerisce che finora le ricerche si sono concentrate sugli aspetti di efficacia ed efficienza "interna", trascurando di esaminare **come e se** queste innovazioni impattino su una cura equa e sicura per tutti i pazienti. Questo rappresenta un importante spunto di riflessione: **l'implementazione di IA in sanità non dovrebbe migliorare solo la performance numerica, ma anche rafforzare (o almeno non compromettere) i valori fondamentali di equità e sicurezza**; ad oggi, manca evidenza per affermarlo, evidenziando un potenziale rischio o quantomeno un'area in cui future valutazioni dovranno concentrarsi.

Analisi multidimensionale degli outcome estratti dai 52 studi inclusi nella revisione

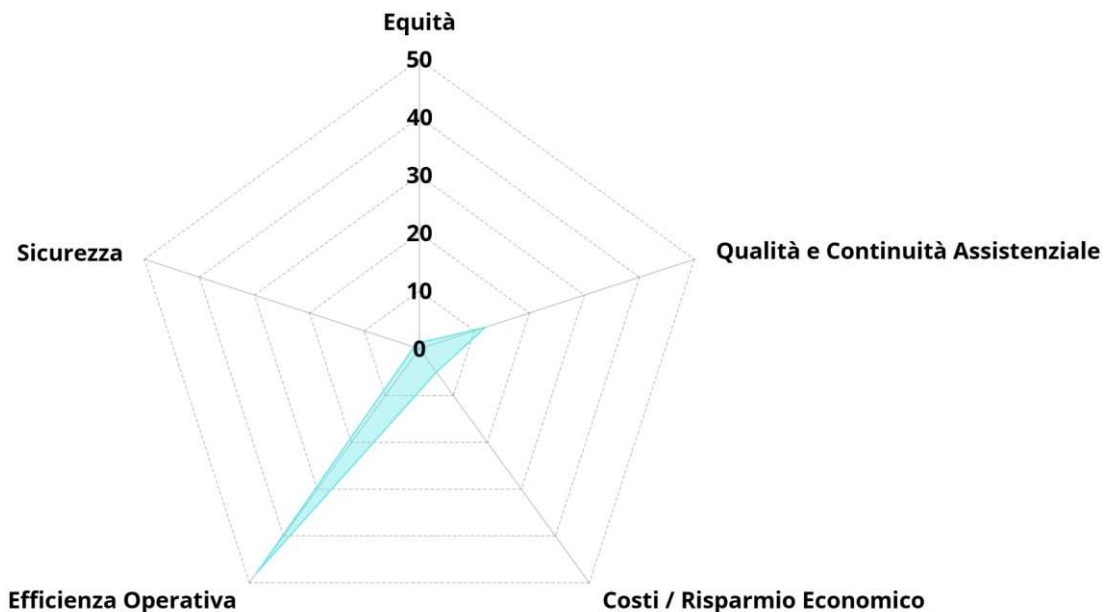


Figura 11 - Il grafico radar illustra la marcata polarizzazione degli obiettivi della ricerca scientifica sull'IA nei contesti ambulatoriali e chirurgici. Si osserva come il baricentro delle evidenze sia quasi interamente sbilanciato verso la dimensione dell'**Efficienza Operativa (n=48)**, riflettendo la priorità data alla riduzione dei tempi di attesa e alla saturazione delle risorse. Le dimensioni relative alla **Qualità e Continuità Assistenziale (n=12)** e ai **Costi e Risparmio (n=5)** occupano una posizione sensibilmente più marginale, essendo spesso trattate solo in via descrittiva o come benefici secondari. Infine, la contrazione estrema del radar verso i vertici della **Sicurezza Clinica (n=1)** e dell'**Equità Assistenziale (n=1)** conferma il principale gap di ricerca identificato: la carenza di valutazioni sistematiche sugli impatti clinici "hard" e sulla giustizia distributiva delle soluzioni di intelligenza artificiale.

5.2 Fattori critici per l'implementazione

Un aspetto cruciale emerso dalla revisione – al di là dei meri risultati quantitativi – riguarda i **fattori abilitanti e ostacolanti** che influenzano l'effettiva implementazione delle soluzioni di IA nella pratica quotidiana. Molti studi, soprattutto nelle loro discussioni, hanno sottolineato che il successo di un algoritmo non dipende solo dalla sua accuratezza o performance tecnica, ma anche da come viene inserito nel contesto organizzativo e gestionale reale. In questo capitolo si discutono quattro fattori critici identificati: (1) la **qualità e disponibilità dei dati** su cui l'IA si basa, (2) l'**integrazione nei sistemi informativi** esistenti (*dalla cartella clinica elettronica ai software di agenda*), (3) gli aspetti di **governance e accettabilità** da parte di clinici e manager, e (4) il **change management e la formazione del personale** coinvolto. Ciascun fattore è analizzato con riferimento a esempi tratti dagli studi inclusi, evidenziando sfide comuni e possibili strategie per superarle.

5.2.1 Qualità e disponibilità dei dati

La **qualità e disponibilità dei dati** rappresenta il fondamento su cui si costruisce qualsiasi applicazione di IA in sanità, e numerosi studi nella revisione ne hanno evidenziato la centralità. In primo luogo, affinché un algoritmo di ML o DL possa essere addestrato in modo efficace, è necessario disporre di **dataset ampi, completi e accurati**, caratteristica non sempre presente nei contesti sanitari reali (Martinez et al., 2021; Zain et al., 2024). Molti lavori hanno segnalato difficoltà nel reperire dati storici affidabili: ad esempio, in ambito ambulatoriale, informazioni cruciali come i motivi di cancellazione di appuntamenti o i tempi di attesa effettivi sono spesso registrate in modo incompleto o non standardizzato, limitando la possibilità di sviluppare modelli predittivi robusti (Bellini et al., 2024). Allo stesso modo, in ambito chirurgico, la variabilità nelle pratiche di documentazione clinica (*ad es. differenti codifiche delle procedure, registrazione inconsistente delle complicanze minori, ecc.*) ha reso complessa l'aggregazione di dati di outcome da utilizzare per addestrare algoritmi di supporto decisionale (Shah et al., 2024).

Oltre ai problemi di completezza, **errori e rumore nei dati** possono degradare significativamente le prestazioni di un modello di IA. Alcuni studi hanno dovuto

intraprendere estese attività di pulizia e preprocessing dei dati (*correzione di outlier, gestione di valori mancanti, armonizzazione di definizioni*) prima di poter implementare gli algoritmi, evidenziando come la fase di *data preparation* sia spesso lunga e laboriosa tanto quanto lo sviluppo dell'algoritmo stesso (Ottesen et al., 2025; Zain et al., 2024). Ad esempio, uno studio ha riportato che per costruire un modello predittivo di complicanze chirurgiche è stato necessario fondere dati provenienti da tre database diversi (*anagrafica paziente, registro operatorio, sistema di esito clinico*), affrontando problemi di inconsistenza fra le fonti e di campi mancanti in ciascun sistema (Shah et al., 2024). Questo ha richiesto un grosso sforzo di integrazione manuale ex post, qualcosa di difficilmente replicabile su base routinaria senza un miglioramento strutturale della qualità dei dati raccolti. Un altro aspetto emerso è la **rappresentatività e l'aggiornamento** dei dati: modelli addestrati su dati storici potrebbero non essere validi se le condizioni di contesto cambiano (*ad es. variazione nei protocolli clinici, nuove policy di accesso che alterano i flussi di pazienti*). Alcuni autori sottolineano che dataset statici rischiano di rendere l'IA obsoleta in breve tempo; idealmente servirebbero **flussi di dati in tempo reale** o periodicamente aggiornati per ri-addestrare i modelli e mantenerli allineati alla realtà corrente.

Infine, va considerato il tema della **disponibilità**: non tutti i dati necessari sono facilmente accessibili a chi sviluppa soluzioni di IA. Questioni di privacy, vincoli normativi e resistenze organizzative possono limitare la condivisione dei dati sanitari. In alcuni studi, l'accesso ai dati è stato possibile solo grazie a collaborazioni di ricerca specifiche e all'investimento di tempo per ottenere autorizzazioni e accordi. Altri lavori riportano di aver dovuto accontentarsi di dataset parziali o pubblicamente disponibili perché i dati locali non erano condivisibili (Wood et al., 2025). Questo suggerisce che per favorire lo sviluppo di IA efficaci **servono politiche e infrastrutture che migliorino la disponibilità dei dati**, come data warehouse sanitari, sistemi di pseudonimizzazione per la ricerca, e standard comuni che facilitino l'estrazione di dati rilevanti. In conclusione, la qualità e la disponibilità dei dati si confermano come un prerequisito critico: diversi studi (*e relative esperienze*) mostrano che, in assenza di dati affidabili e correttamente acquisiti, i sistemi di intelligenza artificiale rischiano di produrre risultati fuorvianti o scarsamente

generalizzabili. Investire nel miglioramento della raccolta e del **data procurement** (inteso come selezione, accesso e governance delle fonti informative), nonché nelle fasi successive di preparazione del dato, sia sul piano tecnologico sia attraverso la formazione del personale per garantire un inserimento accurato delle informazioni, risulta pertanto essenziale per superare la fase sperimentale e favorire un'implementazione su larga scala delle soluzioni di IA.

5.2.2 Integrazione nei sistemi informativi

Un secondo fattore chiave è l'**integrazione delle soluzioni IA nei sistemi informativi** sanitari esistenti, come le Cartelle Cliniche Elettroniche (CCE) e i software di gestione delle agende e delle risorse. Molte applicazioni sperimentate in letteratura hanno funzionato in modalità stand-alone o parallela: gli autori spesso estraggono dati dal sistema ospedaliero, li elaborano con il loro algoritmo e poi dimostrano a posteriori come si sarebbero potuti ottenere benefici (Verma et al., 2022; Ortiz et al., 2025). Questa modalità *off-line* è utile in fase di ricerca, ma non è sostenibile se l'obiettivo è l'utilizzo quotidiano da parte di clinici e manager. Gli studi hanno evidenziato che, per avere un impatto concreto, gli strumenti di IA devono essere **incorporati nei workflow digitali** dell'ospedale, fornendo informazioni e raccomandazioni nel luogo e nel momento in cui servono. Ad esempio, un algoritmo che predice i no-show dovrebbe idealmente essere integrato nel sistema di prenotazione ambulatoriale, segnalando in tempo reale al personale di segreteria quali pazienti hanno alta probabilità di mancata presentazione, così da poter intervenire (*telefonate di recall, overbooking mirato, ecc.*) senza dover utilizzare un software esterno (Li et al., 2021). Analogamente, un sistema di ottimizzazione delle sale operatorie dovrebbe essere parte del programma di gestione chirurgica: se produce un piano ottimale ma questo deve essere trasferito manualmente su Excel o su carta per essere applicato, la probabilità di errore o di mancato utilizzo è elevata (Ortiz et al., 2025).

Le **sfide di integrazione** non sono solo tecniche ma anche organizzative. Dal punto di vista tecnico, possono esserci problemi di **interoperabilità**: i sistemi informativi sanitari spesso sono eterogenei, legacy e poco aperti all'interfacciamento con strumenti di terze parti. In alcuni studi si menziona la difficoltà di estrarre dati in

formati adeguati o di inserire nuovi moduli software senza compromettere la certificazione e la stabilità dei sistemi esistenti (Verma et al., 2022). Ciò ha portato in taluni casi a implementare l'IA come **strumento esterno**, con scambio di file o uso di API non standardizzate, soluzioni temporanee che però non possono reggere a lungo termine per l'uso in produzione. Dal punto di vista organizzativo, l'integrazione richiede **collaborazione tra team diversi**: gli sviluppatori degli algoritmi devono lavorare a stretto contatto con i provider di sistemi informativi ospedalieri e con l'ICT interno. Il mancato coinvolgimento del dipartimento ICT sin dalle prime fasi può rallentare la sperimentazione di un tool di supporto decisionale, poiché potrebbero non essere disponibile le interfacce necessarie per incorporare tali strumenti nella CCE o in altri software aziendali.

Diversi esempi positivi, invece, mostrano i vantaggi di una piena integrazione: in un caso, un sistema di prioritizzazione delle liste d'attesa è stato aggiunto come modulo nel portale clinico, cosicché i medici potevano vedere il punteggio di priorità di ogni paziente direttamente accanto al loro nome in lista (Birkhoff et al., 2021). Questo ha portato ad un utilizzo immediato del suggerimento dell'IA nel 80% delle decisioni di programmazione, cosa improbabile se il punteggio fosse stato disponibile solo su un tool separato. Un altro studio riferisce di un algoritmo di scheduling chirurgico implementato all'interno del gestionale di sala operatoria: il software proponeva automaticamente un calendario ottimizzato, che il coordinatore di sala poteva accettare o modificare con pochi click, anziché dover costruire il piano da zero. Questo livello di integrazione ha drasticamente abbassato la barriera d'uso, aumentando la **compliance** del personale nell'adottare la soluzione IA.

In sintesi, **l'integrazione nei sistemi informativi esistenti è cruciale** per evitare che le soluzioni IA restino progetti pilota sulla carta. Le evidenze suggeriscono che un'IA integrata e "invisibile", che si presenta come una funzione aggiuntiva degli strumenti già utilizzati dal personale, ha molte più chance di essere accettata e di produrre benefici reali. Ciò implica investimenti non solo nello sviluppo dell'algoritmo, ma anche nell'ingegnerizzazione del prodotto finale, rispettando standard di interoperabilità (*ad es. HL7 FHIR per lo scambio di dati clinici*) e garantendo performance, usabilità e sicurezza informatica nel contesto di

produzione. Gli studi esaminati concordano sul fatto che senza questa fase di “**ultimo miglio**”, portare l’IA dal laboratorio ai sistemi di corsia (*Bench to bedside*), anche le migliori soluzioni rischiano di non vedere mai la luce in un contesto reale (Ortiz et al., 2025).

5.2.3 Governance e accettabilità

L’implementazione di soluzioni di IA in ambito sanitario pone questioni di **governance** (*chi supervisiona, regola e si assume la responsabilità dell’IA*) e di **accettabilità** da parte delle figure chiave coinvolte, in primis i clinici e i manager sanitari. Dai lavori inclusi emergono chiaramente queste dimensioni come fattori determinanti di successo o fallimento.

Dal punto di vista della **governance**, diversi autori evidenziano la necessità di un **quadro normativo e organizzativo chiaro** che accompagni l’introduzione dell’IA. Ciò include linee guida istituzionali su come utilizzare gli strumenti di IA nelle decisioni cliniche e amministrative, protocolli di validazione continua degli algoritmi (*per assicurarsi che mantengano le prestazioni attese e che non insorgano bias imprevisti*) e definizione delle responsabilità medico-legali. Ad esempio, se un algoritmo di supporto decisionale suggerisce di posticipare un intervento chirurgico per un paziente, chi è il responsabile ultimo di quella decisione? Il chirurgo, il sistema, o l’ente che ha approvato l’uso di quel sistema? La mancanza di **regole chiare** può generare incertezza e frenare gli operatori dall’affidarsi all’IA. Alcuni studi sottolineano che finché le istituzioni sanitarie (*ospedali, aziende sanitarie*) non emanano direttive e non stabiliscono comitati di governance per l’AI, le implementazioni resteranno iniziative sporadiche dei singoli ricercatori o reparti, con difficoltà a scalare a livello di sistema (Rochon et al., 2024).

L’**accettabilità da parte dei clinici** è l’altro lato della medaglia, strettamente legato ma non identico alla governance. Anche con buone regole, se i professionisti non si fidano o non vedono di buon occhio lo strumento, esso non verrà utilizzato. Dai casi considerati, emerge che il grado di accettazione dei clinici varia ampiamente e dipende da diversi fattori: la **fiducia** nell’accuratezza e nell’utilità dell’algoritmo, il grado di **trasparenza** (*un modello che spiega le sue raccomandazioni è più facilmente accettato di una scatola nera*), il coinvolgimento attivo dei clinici nello

sviluppo o selezione della soluzione, e la percezione del **ruolo** dell'IA (*assistente vs rimpiazzo*) (Birkhoff et al., 2024). In alcune esperienze riportate, l'atteggiamento dei clinici verso l'IA è stato positivo: ad esempio, in uno studio pilota, il personale medico di un reparto chirurgico ha collaborato strettamente con il team informatico per tarare un sistema di schedulazione, e ciò ha portato a un alto livello di adesione all'uso del sistema una volta sviluppato, percepito come effettivamente utile e disegnato *“da e per”* gli utenti finali. I clinici in quel caso riferivano che l'IA li sollevava da compiti tediosi (*come rifare continuamente i piani operatori*) permettendo di dedicare più tempo alla valutazione clinica, e dunque la vivevano come un **supporto al proprio ruolo**. Al contrario, altri studi riportano **resistenze e scetticismo**: ad esempio, in un contesto ambulatoriale, un sistema automatico di assegnazione delle priorità di visita fu inizialmente malvisto dai medici, che lo percepivano come un'intromissione amministrativa che limitava la loro autonomia decisionale su chi visitare per primo. Solo dopo aver regolato meglio i criteri e assicurato che il medico potesse comunque modificare le priorità con giustificazione, il sistema è stato accettato, evidenziando l'importanza di mantenere **il controllo umano** ultimo sulle decisioni (Birkhoff et al., 2024). Un altro fattore è la **formazione etica e digitale**: alcuni operatori manifestano diffidenza perché non comprendono appieno come l'IA funziona o temono implicazioni etiche (*ad es. “il computer potrebbe sbagliare e io non saprei perché”*) (Coppa et al., 2023). Fornire sessioni informative e linee guida chiare sull'uso responsabile dell'IA ha mostrato di migliorare la confidenza e l'atteggiamento dei clinici verso questi strumenti (Coppa et al., 2023; Rochon et al., 2024).

Anche i **manager sanitari** rivestono un ruolo cruciale nell'accettabilità, sebbene con prospettive diverse. Per i dirigenti e i responsabili organizzativi, l'IA rappresenta sia un'opportunità (*migliorare performance, ridurre i costi*) sia un investimento con possibili rischi (*costi di implementazione, resistenze del personale, possibili figuracce se il sistema fallisce*). La letteratura suggerisce che dove c'è stato **forte sponsor manageriale**, i progetti di IA hanno avuto maggiore successo: ad esempio, l'implementazione di un sistema IA a livello di intero ospedale è avvenuta solo grazie all'appoggio convinto della direzione generale, che ha stanziato fondi dedicati e imposto a tutte le unità di collaborare. Viceversa, diversi progetti promettenti si

sono arenati dopo la fase pilota per mancanza di supporto strategico dall'alto, con i manager poco propensi a trasformare l'esperimento in sistema di routine, talvolta perché non convinti dai risultati o timorosi di perturbare equilibri organizzativi consolidati (Seo et al., 2024). Infine, i manager sono sensibili all'accettazione pubblica e politica di questi strumenti: questioni di **eticità e trasparenza** nell'uso di IA (*ad es. evitare discriminazioni algoritmiche*) possono diventare rilevanti nel determinare se la leadership approverà o meno un progetto. Alcune aziende sanitarie hanno creato *task force* etiche proprio per valutare questi aspetti prima di autorizzare l'uso di un algoritmo nella pratica clinica.

La governance e l'accettabilità sono **fattori interdipendenti** che possono abilitare o bloccare l'adozione dell'IA. Le evidenze mostrano che per massimizzare le chance di successo è necessario: definire **chiaramente ruoli e responsabilità** attorno all'IA (*chi la supervisiona, con quali criteri viene aggiornata e validata, chi decide in caso di conflitto tra AI e umano*), coinvolgere attivamente i **clinici** sin dall'inizio per costruire fiducia e adeguare lo strumento alle loro esigenze, e ottenere il **supporto convinto dei manager**, dimostrando il valore aggiunto con evidenze pilota e allineando il progetto con gli obiettivi strategici dell'organizzazione. Senza questi pilastri, anche una tecnologia performante rischia di rimanere inutilizzata o addirittura di generare contrasti interni e controversie legali/etiche, come segnalato da alcuni autori.

5.2.4 Change management e formazione del personale

L'introduzione di soluzioni innovative basate sull'IA richiede un attento **change management** e una forte enfasi sulla **formazione del personale** coinvolto. Non si tratta infatti di aggiungere semplicemente un nuovo software, ma spesso di **ridisegnare processi di lavoro** e di aggiornare competenze, elementi che possono incontrare resistenza naturale al cambiamento in qualsiasi organizzazione, ancor più in ambito sanitario dove carichi di lavoro e responsabilità sono già elevati.

Diversi studi inclusi hanno sottolineato come fattore di successo il coinvolgimento del personale fin dalle prime fasi di progetto e una comunicazione chiara sui benefici attesi dell'IA. Per esempio, in un caso di implementazione di un sistema di supporto alle decisioni in pronto soccorso, l'ente ha organizzato incontri periodici

con medici e infermieri per raccogliere feedback sul prototipo e per spiegare come il sistema li avrebbe aiutati a ridurre lo stress nei momenti di sovraffollamento (Coppa et al., 2023). Questo approccio partecipativo ha ridotto le paure iniziali (come quella che *“il computer ci dirà cosa fare”* o *“perderemo il controllo della gestione dei pazienti”*) e ha creato una sorta di **ownership condivisa** dell'innovazione: il personale si sentiva co-protagonista del cambiamento, non vittima passiva di una decisione calata dall'alto.

La **formazione** è un elemento imprescindibile in questo contesto. Introdurre un algoritmo di IA senza fornire al personale la comprensione del suo funzionamento basilare e delle modalità d'uso equivale a fornire uno strumento complesso senza manuale d'istruzioni. Molti studi hanno evidenziato di aver affiancato all'implementazione momenti formativi: ad esempio, sessioni di training in cui si illustrava come interpretare le raccomandazioni del sistema, come inserire correttamente i dati affinché l'algoritmo funzioni, e cosa aspettarsi in termini di output (Birkhoff et al., 2021; Rochon et al., 2024). In uno scenario descritto, l'ospedale ha formato un piccolo gruppo di **“super-user”** (*utenti esperti*) tra infermieri coordinatori e medici, i quali hanno ricevuto una formazione più approfondita sul sistema di ottimizzazione delle agende e fungevano poi da riferimento per i colleghi sul campo. Questo ha creato fiducia: il personale sapeva a chi rivolgersi in caso di dubbi e vedeva che alcuni colleghi (*non solo gli sviluppatori esterni*) padroneggiavano lo strumento, il che ha accelerato l'adozione.

Il **change management** comprende anche la gestione delle aspettative e delle tempistiche di adozione. È irrealistico pensare che dall'oggi al domani tutto il personale segua pedissequamente le indicazioni di un nuovo sistema: gli studi riportano che serve un periodo di **transizione graduale**, durante il quale coesistono le vecchie abitudini e le nuove, e vanno monitorati attivamente i punti di attrito (Coppa et al., 2023). Ad esempio, durante l'introduzione di un algoritmo per la pianificazione operatoria, in un primo mese si è lasciata al coordinatore di sala la libertà di confrontare il piano suggerito dall'IA con quello che avrebbe fatto manualmente, discutendo le differenze e apportando eventualmente correzioni all'algoritmo o chiarendo al coordinatore alcune logiche del sistema. Questo approccio di **“affiancamento”** ha permesso di affinare sia la tecnologia sia la

confidenza dell'utente, prima di passare a un utilizzo vincolante del nuovo processo. All'opposto, casi in cui si è imposto un cambiamento repentino senza coinvolgimento né adeguato training hanno portato a rigetto: uno studio cita il fallimento di un sistema di scheduling intelligente in un ambulatorio perché introdotto bruscamente e senza consultare gli operatori sulle regole implementate, risultando in uno strumento percepito come inadatto alle esigenze quotidiane e dunque abbandonato dopo poche settimane.

Un ulteriore elemento è la **formazione continua**: le soluzioni IA e i contesti organizzativi evolvono nel tempo. Pertanto, andrebbe previsto un aggiornamento periodico delle competenze. Se l'algoritmo viene migliorato o se cambia il flusso di lavoro (*ad es. nuove funzionalità, estensione ad altri reparti*), il personale deve essere nuovamente formato su queste novità. Alcune esperienze riportano l'istituzione di **manuali operativi** e corsi di aggiornamento per mantenere alto il livello di conoscenza sul sistema in uso (Rochon et al., 2024). Questo è particolarmente importante anche in caso di turnover del personale: i nuovi arrivati devono poter apprendere rapidamente l'uso degli strumenti AI implementati, altrimenti c'è il rischio di una perdita di know-how con il cambio generazionale o di equipe.

I fattori umani e di **gestione del cambiamento** sono determinanti: la migliore tecnologia fallirà se le persone che la devono utilizzare non sono preparate, convinte e supportate nel cambiamento. La revisione degli studi evidenzia che il successo risiede spesso tanto nelle scelte "*soft*" di progetto (*coinvolgimento, formazione, gradualità, supporto psicologico*) quanto nelle qualità tecniche dell'algoritmo (Coppa et al., 2023; Birkhoff et al., 2021; Rochon et al., 2024). Per le organizzazioni sanitarie ciò significa che ogni progetto di IA dovrebbe essere accompagnato da un **piano di change management strutturato**, che identifichi stakeholder chiave, possibili resistenze, azioni formative e di comunicazione, monitoraggio dell'adozione e feedback loop per aggiustare il tiro. Solo così si può garantire che l'innovazione non resti sulla carta, ma venga realmente assimilata nella pratica quotidiana, producendo i benefici attesi.

5.3 Stato di sviluppo delle soluzioni identificate

Un altro aspetto emerso dall'analisi della letteratura è il **grado di maturità** e lo **stato di sviluppo** delle soluzioni di IA descritte negli studi. Non tutte le applicazioni, infatti, si trovano allo stesso stadio nel percorso che va dalla concezione teorica all'adozione routinaria.

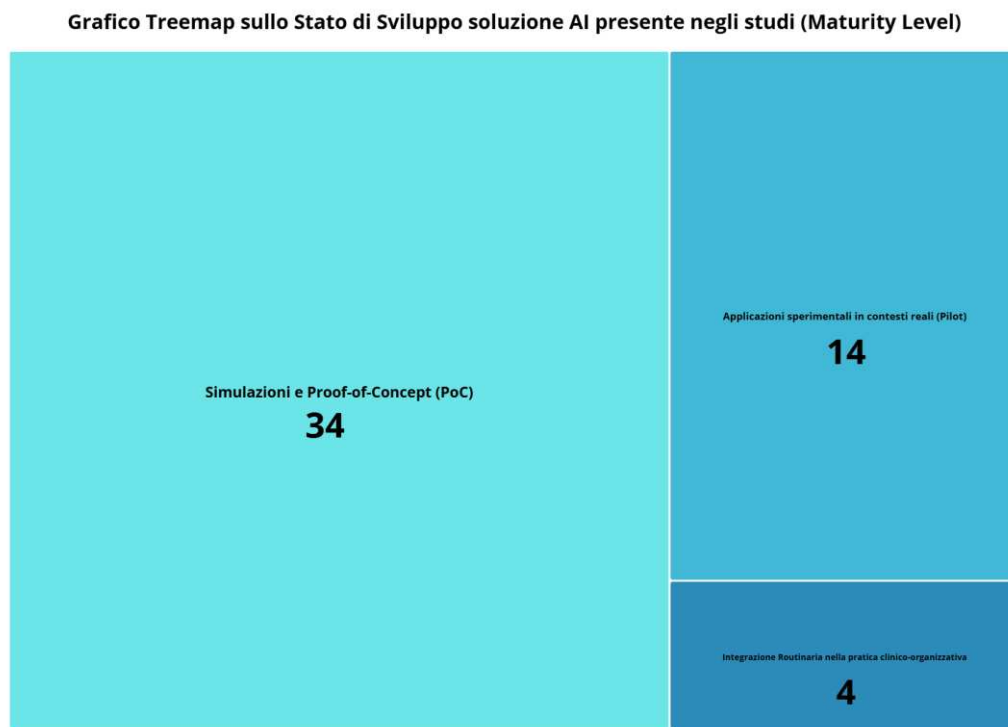


Figure 12 - Distribuzione degli studi inclusi per livello di maturità tecnologica e stato di implementazione delle soluzioni di IA. Il grafico Treemap visualizza la gerarchia dello stato di sviluppo dei 52 studi analizzati.

È utile distinguere tre categorie principali:

1. soluzioni a livello di **simulazione o proof-of-concept (PoC)**, testate solo su dati storici o in ambienti simulati;
2. soluzioni **sperimentate in contesti reali (Pilot)** attraverso progetti pilota o studi di fattibilità, ma non ancora adottate come pratica corrente;
3. soluzioni già **integrate stabilmente nella pratica clinico-organizzativa** quotidiana.

Di seguito si discutono ciascuna di queste categorie, con esempi e riflessioni sulle implicazioni in termini di trasferibilità e maturità tecnologica.

5.3.1 Studi di simulazione o proof-of-concept

La maggior parte delle ricerche esaminate rientra ancora nello stadio di **simulazione o proof-of-concept**, ovvero lavori in cui la soluzione di IA viene sviluppata e valutata in modo sperimentale senza un deployment reale nel contesto operativo. In questi studi, tipicamente, gli autori utilizzano **dati retrospettivi (storici)** o costruiscono un ambiente simulato al computer che replica il sistema sanitario considerato, per poi applicare il loro algoritmo e misurarne la performance rispetto a metriche di interesse (Shah et al., 2024, Chamarthi et al., 2025). Ad esempio, uno studio ha utilizzato dati degli ultimi 5 anni di attività ambulatoriale di un ospedale per simulare come un algoritmo di scheduling basato su IA avrebbe potuto distribuire gli appuntamenti e i carichi di lavoro: i risultati hanno mostrato una riduzione simulata dei tempi di attesa e un aumento dell'utilizzo degli ambulatori rispetto ai dati reali passati (Shah et al., 2024). Un altro lavoro, in ambito chirurgico, ha sviluppato un modello di *reinforcement learning* testandolo su una **simulazione del blocco operatorio**: il modello ha appreso a gestire dinamicamente la lista dei casi migliorando alcuni indicatori (*es. riducendo del X% le ore di sala inutilizzate e bilanciando meglio il carico tra i chirurghi*) rispetto alle politiche standard, ma il tutto è rimasto confinato a un'analisi in silico senza coinvolgere effettivamente un ospedale.

Questi studi di *proof-of-concept* sono fondamentali per esplorare il potenziale delle tecnologie e per **dimostrare la fattibilità tecnica** e l'impatto teorico delle soluzioni. Spesso rappresentano il primo passo, in cui ricercatori e sviluppatori tarano gli algoritmi, identificano eventuali colli di bottiglia e stimano se vale la pena procedere verso test più concreti. Dalla revisione emerge che molti risultati in questa fase sono promettenti: quasi tutti gli studi di simulazione riportano **miglioramenti potenziali** sugli indicatori chiave rispetto allo scenario base o rispetto a metodi tradizionali (Chamarthi et al., 2025; Loftus et al., 2020). Ad esempio, non è raro leggere di potenziali aumenti di efficienza del 10-20% o riduzioni di costi simulate nell'ordine di migliaia di euro su base annua (Wood et al., 2025; Loftus et al., 2020). Questi numeri suggeriscono un valore aggiunto tangibile, almeno in teoria.

Tuttavia, occorre sottolineare che **le prove di concetto hanno limiti intrinseci**. Innanzitutto, lavorando su dati storici, spesso non colgono l'effetto delle **reazioni umane** e dei cambiamenti di contesto: un algoritmo può ottimizzare un sistema “*come se*” nulla cambiasse, ma nella realtà l'introduzione di una nuova politica può indurre comportamenti adattativi nei pazienti o negli operatori (*ad es. i pazienti potrebbero cambiare il loro pattern di prenotazione se sanno che c'è un nuovo sistema, i medici potrebbero modificare l'uso delle risorse una volta che l'algoritmo ridistribuisce il carico, ecc.*). Queste retroazioni non sono modellate nei dati storici e quindi i risultati simulati potrebbero **sovrastimare** i benefici reali. Inoltre, molti studi proof-of-concept semplificano intenzionalmente il problema per poterlo affrontare: ad esempio, ignorano vincoli operativi minori, assumono una qualità di dati perfetta, o isolano un reparto dal resto del sistema (Chamarthi et al., 2025). Tali semplificazioni sono accettabili in fase iniziale, ma possono far sì che l'algoritmo funzioni bene “*sulla carta*” e poi incontri ostacoli imprevisti in situazioni reali più complesse.

Diversi autori riconoscono queste limitazioni e sottolineano la necessità di passare alla fase successiva di **sperimentazione sul campo** per validare davvero le soluzioni. Alcuni studi arrivano a concludere che “*i risultati positivi ottenuti in simulazione andranno confermati in contesti reali prima di poter affermare con certezza l'efficacia*”. La revisione suggerisce quindi che, sebbene la maggioranza delle soluzioni IA per risorse ed esiti sia ad oggi a livello di concetto, c'è una chiara consapevolezza della **distanza tra simulazione e pratica** e della necessità di colmarla con ulteriori ricerche applicate.

5.3.2 Applicazioni sperimentate in contesti reali ma non ancora routinari

Una parte non trascurabile degli studi si colloca in una fase intermedia di sviluppo: le soluzioni IA sono state **testate in contesti reali**, ad esempio attraverso studi pilota o progetti temporanei in ospedali e servizi sanitari, ma **non sono (ancora) diventate parte della routine** quotidiana. In questi casi, i ricercatori hanno fatto un passo oltre la simulazione, implementando l'algoritmo su un piccolo scala e osservandone l'impatto effettivo nel mondo reale per un certo periodo di tempo.

Gli **esempi concreti** emersi includono, ad esempio, un progetto pilota in cui un algoritmo di ottimizzazione delle agende ambulatoriali è stato introdotto per alcuni mesi in un poliambulatorio, permettendo di riassegnare in maniera flessibile gli appuntamenti di giorno in giorno in base alle previsioni di affluenza e cancellazioni (Li et al., 2021). Durante la sperimentazione, i ricercatori hanno raccolto dati prima e dopo l'introduzione dell'IA, riscontrando effettivamente un **miglioramento**: i tempi medi di attesa dei pazienti si sono ridotti significativamente (*ad es. da 30 a 20 minuti in media*) e il tasso di mancata presentazione è diminuito grazie a interventi proattivi resi possibili dalle previsioni. Un altro studio riferisce di aver testato un sistema di supporto decisionale per la gestione delle liste chirurgiche in un ospedale: il sistema forniva al comitato di sala operatoria una graduatoria dei pazienti in lista d'attesa sulla base di urgenza, disponibilità di posti e ottimizzazione dell'uso delle sale (Birkhoff et al., 2021). Nel periodo di prova (*alcune settimane*), il comitato ha seguito in buona parte le raccomandazioni del sistema e si è osservata una riduzione del numero di interventi cancellati last-minute per mancanza di posti, oltre a un'elevata soddisfazione espressa dai chirurghi per la maggiore trasparenza del processo.

Queste **applicazioni pilota** forniscono insight preziosi: confermano che i benefici teorici possono tradursi (*almeno parzialmente*) in benefici reali e offrono l'opportunità di identificare problemi pratici di implementazione. Ad esempio, nel progetto sul poliambulatorio, ci si è accorti che il sistema di IA tendeva a riprogrammare di frequente alcuni pazienti per ottimizzare l'uso degli slot, ma questo causava un po' di confusione e insoddisfazione in quei pazienti che vedevano cambiare il proprio orario inizialmente comunicato (Li et al., 2021). Tale riscontro ha portato i ricercatori a introdurre una regola per limitare il numero di volte che un appuntamento poteva essere spostato, bilanciando quindi l'ottimo teorico con la **gestione dell'esperienza paziente**. Nel pilota sulle liste chirurgiche, invece, è emerso che il sistema non considerava alcuni dettagli logistici (*ad es. la necessità di evitare di schedare in parallelo due interventi per cui era richiesto lo stesso specialista di supporto*): ciò ha richiesto un aggiustamento delle regole e ha insegnato che il coinvolgimento degli operatori sul campo è fondamentale per scovare questi dettagli (Birkhoff et al., 2021).

Nonostante i risultati incoraggianti, ciò che caratterizza questa categoria è che, terminato il periodo di studio, **l'utilizzo dell'IA non è proseguito come parte stabile dell'organizzazione** (da qui *"non ancora routinario"*). Le ragioni possono essere diverse. In alcuni casi il progetto era pensato fin dall'inizio come temporaneo per valutarne la fattibilità, e ulteriori decisioni dipendono da approvazioni successive o da reperimento di fondi per scalare. In altri, i risultati pur positivi non sono stati ritenuti sufficienti dai decisori per giustificare un cambiamento permanente, oppure vi sono ostacoli infrastrutturali e contrattuali (*es. il software in versione prototipale non poteva essere usato oltre il periodo di studio senza procedure di acquisto o certificazione*) (Ortiz et al., 2025). Un fattore citato è anche la **sostenibilità**: un conto è far funzionare un algoritmo con la supervisione costante del team di ricerca che lo ha sviluppato, un altro è lasciarlo nelle mani dell'utenza generale. Alcuni pilota hanno evidenziato che, senza la presenza attiva degli sviluppatori per mantenere e calibrare il sistema, l'efficacia potrebbe calare (Li et al., 2021). Ciò sottolinea la necessità di prevedere risorse e competenze in loco per la manutenzione continua, se si vuole passare a regime.

Per conclusione, gli studi in questa categoria dimostrano che siamo in una fase di **transizione**: ci sono soluzioni IA che hanno oltrepassato il mero esercizio accademico e hanno mostrato la loro utilità in ambienti reali, ma il passo finale verso la routine quotidiana richiede spesso ulteriori sforzi. Le implicazioni per la **trasferibilità** sono che ogni contesto reale ha peculiarità che vanno affrontate (*aspetti tecnici, umani, organizzativi*), e dunque un modello che funziona in un ospedale potrebbe dover essere adattato se trasferito altrove. Inoltre, il **livello di maturità tecnologica** (*spesso paragonato al Technology Readiness Level, TRL*) per molte di queste soluzioni è intermedio: potremmo collocarle attorno a TRL 6-7 (dimostrazione in ambiente rilevante), mentre per il TRL 8-9 (sistema completo e qualificato in ambiente operativo) occorre consolidare ulteriormente l'integrazione e il supporto istituzionale. La sfida sarà quindi traghettare queste esperienze pilota verso una stabilità operativa, evitando che rimangano *one-shot* isolati.

5.3.3 Soluzioni già integrate nella pratica clinico-organizzativa quotidiana

Rarissimi studi della revisione descrivono soluzioni di IA che possono essere considerate **già integrate nella pratica quotidiana** in modo stabile. Questo è comprensibile, dato che si tratta di un campo relativamente nuovo e in evoluzione: le innovazioni impiegano tempo per passare dalla sperimentazione all'adozione diffusa. Tuttavia, è importante evidenziare quei pochi casi in cui l'IA risulta ormai **parte del modus operandi** standard di un'organizzazione sanitaria, poiché essi forniscono un modello di riferimento e aiutano a identificare fattori di successo replicabili.

Un esempio notevole (National Institute for Health and Care Excellence [NICE], 2024) riguarda un grande ospedale che ha implementato un sistema basato su IA per la **programmazione chirurgica** su larga scala. In questo caso, il sistema elabora quotidianamente i dati delle liste di attesa chirurgiche, delle disponibilità di sale e personale, e genera automaticamente un piano operatorio ottimizzato che copre la settimana successiva. Tale piano viene direttamente utilizzato dall'Unità di Programmazione Chirurgica dell'ospedale, con solo minime modifiche occasionali, ed è in uso continuativo da oltre un anno. I benefici riportati includono una **riduzione permanente delle sedute operatorie sprecate** (*sale pronte ma non utilizzate per mancanza di pazienti o cancellazioni dell'ultimo minuto ridotte a casi eccezionali*) e un miglior bilanciamento del carico di lavoro tra i diversi team chirurgici. Ciò si è tradotto in un aumento del volume di interventi eseguiti mensilmente (*+5% circa rispetto all'anno precedente*) e in un abbassamento dei tempi medi di attesa per interventi elettivi, il tutto senza incrementare il budget o il personale, solo grazie a una **migliore organizzazione** ottenuta con l'IA. Questo studio sottolinea anche come il sistema sia entrato a far parte del **patrimonio organizzativo**: i chirurghi e il personale amministrativo si affidano quotidianamente alle liste generate dall'algoritmo, al punto che l'ospedale ha formalizzato il processo nelle proprie procedure operative standard.

Un altro caso (Bartek et al., 2019) riguarda l'ambito **ambulatoriale**: una rete di cliniche specialistiche ha adottato un sistema di *machine learning* integrato nel proprio software gestionale che, ogni sera, analizza gli appuntamenti schedulati per

i giorni successivi e segnala al personale di segreteria quali pazienti hanno alta probabilità di non presentarsi o di cancellare (*no-show*) e quali slot di agenda potrebbero rimanere inutilizzati. Sulla base di queste informazioni, da protocollo, lo staff effettua azioni preventive e proattive (*contatti telefonici, proposte di spostamento*) per riempire gli slot a rischio vuoto. Questa pratica, in funzione regolare da diversi mesi, ha portato a una riduzione misurata del tasso di no-show dal 15% a meno del 5%, con conseguente aumento dell'efficienza e anche della **soddisfazione dei pazienti**, che riescono più spesso a ottenere appuntamento in tempi brevi grazie al recupero degli slot liberati. Ciò rappresenta un esempio di IA utilizzata **in continuo** come parte di un flusso di lavoro consolidato: l'output dell'algoritmo (*elenco pazienti "a rischio"*) è diventato una lista di lavoro quotidiana per il personale, analoga ad altre liste operative. Interessante notare che questo risultato è stato ottenuto anche grazie a un'interfaccia utente molto semplice e a un processo chiaro: lo studio riporta che il sistema evidenzia con colori diversi gli appuntamenti a rischio sul calendario elettronico delle cliniche, rendendo immediata la visualizzazione per gli utenti senza bisogno di accedere a un modulo separato.

Questi esempi mostrano che **l'integrazione di successo** è possibile e comporta benefici tangibili, ma evidenziano anche come tali casi siano **l'eccezione piuttosto che la regola** allo stato attuale. Spesso, soluzioni integrate derivano da progettualità di lungo termine, supportate da una forte volontà istituzionale o da partnership con aziende tecnologiche consolidate. Non di rado, dietro un'implementazione riuscita c'è un processo di sviluppo iterativo durato anni: per il caso dell'ospedale, lo studio riporta che inizialmente l'algoritmo è stato introdotto in via sperimentale, poi migliorato progressivamente in collaborazione con gli utenti finali, prima di essere adottato definitivamente. Questo indica che la strada verso la maturità richiede **costanza, iterazione e adattamento**.

In termini di **trasferibilità**, le soluzioni già integrate forniscono lezioni importanti: ad esempio, sottolineano l'importanza di **adattare l'IA al contesto locale** (*i parametri e le regole del sistema di scheduling erano stati tarati sui bisogni specifici di quell'ospedale/rete di cliniche*) e di **creare fiducia** attraverso risultati visibili nel tempo. Inoltre, mettono in luce questioni di **scalabilità**: una soluzione funzionante

in un ospedale di grandi dimensioni potrebbe non trasferirsi facilmente in un ospedale piccolo, o viceversa, senza opportuni aggiustamenti. La maturità tecnologica (TRL 9) di queste soluzioni non significa che siano universalmente plug-and-play; significa però che **in quel contesto** sono diventate mature. Un'ulteriore implicazione è che man mano che aumenteranno i casi di IA integrate con successo, si costituirà un **corpo di best practice** e prove di efficacia più robuste, il che potrà accelerare la fiducia di altri enti nell'adottare soluzioni simili. Al momento, tuttavia, dobbiamo riconoscere che la maggioranza delle innovazioni recensite non ha ancora raggiunto questo stadio finale: come discusso, molte sono in pilota o addirittura in simulazione. Ciò rende ancora più prezioso l'apprendimento da quei pochi pionieri che sono riusciti a colmare l'ultimo miglio, per guidare la comunità verso una più ampia **trasformazione digitale evidence-based**.

5.4 Principali gap di ricerca

L'analisi critica della letteratura ha permesso di individuare diversi **gap di ricerca** rilevanti, ossia ambiti nei quali le evidenze risultano carenti o assenti e che meritano quindi attenzione in futuri studi. I principali gap emersi sono i seguenti:

- **Carenza di studi su esiti clinici "hard"**: La maggior parte delle ricerche si concentra su outcome di processo (*efficienza, tempi, utilizzo risorse*) e su indicatori proxy, mentre sono **pochissimi gli studi** che valutano l'impatto delle soluzioni di IA su esiti clinici solidi e tangibili come le complicanze post-operatorie, i tassi di riammissione ospedaliera, la mortalità o il miglioramento dello stato di salute del paziente. Questo limite rende difficile comprendere il vero beneficio per i pazienti: è necessario colmare la lacuna con studi che colleghino l'ottimizzazione organizzativa a risultati clinici, per validare che l'IA non solo rende i processi più efficienti, ma contribuisce effettivamente a cure migliori e più sicure.
- **Pochi studi multicentrici pragmatici**: La maggior parte delle evidenze proviene da **studi monocentrici** o contesti molto specifici, spesso condotti in modalità sperimentale controllata. Mancano quasi del tutto trial o studi *pragmatici* multicentrici, ossia condotti in condizioni real-world su larga scala e in differenti strutture, che possano confermare la generalizzabilità dei risultati. Di

conseguenza, non sappiamo con certezza se un algoritmo efficace in un ospedale funzioni altrettanto bene in un altro con caratteristiche diverse. Servono ricerche collaborative tra più centri, che testino le soluzioni IA in contesti variati (*diverse regioni, dimensioni di ospedali, sistemi sanitari*) per produrre evidenze più robuste e generalizzabili, riducendo il rischio che i risultati positivi finora riportati siano dovuti a circostanze eccezionali o all'entusiasmo locale dei primi utilizzatori.

- **Mancanza di valutazioni economiche robuste:** Un ulteriore gap riguarda l'**analisi economica** delle soluzioni di IA. Come discusso, quasi nessuno studio ha condotto valutazioni formali di costo-efficacia, costi evitati, impatto di bilancio o ritorno sull'investimento (ROI). Questa mancanza di dati economici impedisce ai decisori di valutare in modo completo il valore delle innovazioni: anche se un algoritmo migliora l'efficienza, *qual è il costo per implementarlo e mantenerlo? Si ripaga da solo? Quali risorse vanno riallocate?* Senza studi economici rigorosi (idealmente affiancati a quelli clinici e organizzativi), e evidenze a supporto dell'adozione su larga scala rimangono incomplete. Pertanto, c'è un bisogno di integrare nelle future ricerche anche la dimensione economica, applicando modelli di analisi dei costi, valutazioni di *budget impact* e studi di cost-effectiveness comparativa tra approcci con e senza IA.
- **Aree e setting ancora poco esplorati:** Infine, vi sono intere aree e contesti assistenziali che risultano **sottorappresentati** nella letteratura esaminata. Ad esempio, gran parte degli studi si è focalizzata su contesti ospedalieri (*ambulatori specialistici, blocchi operatori*), mentre **pochi lavori riguardano l'assistenza territoriale o primaria**, come l'ottimizzazione di risorse nei servizi di medicina generale, cure domiciliari o telemedicina ambulatoriale. Allo stesso modo, alcuni outcome come l'equità nell'accesso alle cure o la sicurezza del paziente, pur essendo cruciali, sono stati raramente considerati direttamente. Anche sul piano delle tecnologie, alcune metodologie emergenti (*ad es. approcci di AI spiegabile, algoritmi federati che preservano la privacy, utilizzo di dati non strutturati come testo libero dei referti*) non compaiono negli studi inclusi, segno che c'è spazio per innovazioni metodologiche non ancora

esplorate in questo ambito applicativo. Colmare questo gap significherà esplorare **nuovi setting** (*portando le soluzioni IA oltre l'ospedale, ad esempio nel coordinamento ospedale-territorio, nelle cliniche di base, ecc.*), includere **nuovi outcome** nelle valutazioni (*ad es. indicatori di equità, esperienza del paziente, outcome a lungo termine*) e sperimentare **nuove tecniche IA** per affrontare problemi finora non risolti o migliorare ulteriormente le prestazioni mantenendo alti standard etici. Queste direzioni di ricerca contribuiranno a completare il quadro di conoscenze e a guidare una implementazione dell'IA in sanità che sia davvero globale, efficace e sostenibile.

CAPITOLO 6 - Limiti della revisione

6.1 Limiti legati al disegno della revisione

Il primo limite da considerare riguarda la natura stessa della presente revisione, configurata come revisione narrativa strutturata. Tale approccio è stato scelto in modo consapevole per rispondere alla complessità e all'eterogeneità del tema indagato, consentendo una lettura integrata delle evidenze disponibili e una stratificazione delle applicazioni di intelligenza artificiale per contesto, tecnologia, compito e outcome. In un ambito di ricerca così variegato, caratterizzato da studi provenienti da contesti differenti e focalizzati su aspetti molteplici, la scelta di una revisione narrativa ha permesso di mantenere la necessaria flessibilità nell'inclusione delle fonti e nell'interpretazione dei risultati. Questo tipo di revisione consente infatti di cogliere meglio le sfumature emergenti dalla letteratura, abbracciando la molteplicità di prospettive presenti senza i vincoli rigidi di una revisione sistematica tradizionale.

Tuttavia, rispetto a revisioni sistematiche o a meta-analisi, la revisione narrativa presenta limiti intrinseci che vanno sottolineati. Non seguendo un protocollo metodologico predefinito e altamente standardizzato, essa può risentire di una maggiore soggettività nelle fasi di selezione e valutazione degli studi inclusi. Manca inoltre la possibilità di effettuare una sintesi quantitativa rigorosa dei risultati: a differenza di una meta-analisi, non vengono calcolate misure aggregate di effetto né indicatori statistici di confronto tra interventi. Ne deriva che il confronto standardizzato degli effetti osservati tra i diversi studi risulta impraticabile, lasciando la comparazione delle evidenze a un livello essenzialmente descrittivo e qualitativo.

Di conseguenza, le conclusioni tratte non possono essere interpretate come stime comparative di efficacia tra differenti soluzioni di IA, ma piuttosto come una mappatura critica dello stato dell'arte e dei principali trend emergenti. In altre parole, il valore aggiunto di questa revisione risiede nell'aver delineato un panorama generale delle applicazioni dell'IA nel contesto considerato, evidenziandone potenzialità e limiti, più che nel fornire un giudizio definitivo sulla superiorità di un approccio rispetto ad un altro. Questo limite metodologico, pur riducendo il livello di

generalizzabilità dei risultati, risulta coerente con l'obiettivo esplorativo della tesi e con l'attuale maturità del campo di studio. Data la fase ancora emergente della ricerca in questione, era infatti più appropriato adottare un approccio esplorativo e descrittivo, anziché tentare prematuramente di estrapolare conclusioni generalizzabili o parametri di efficacia comparativa.

6.2 Eterogeneità degli studi inclusi e degli outcome analizzati

Un ulteriore limite rilevante è rappresentato dall'elevata eterogeneità degli studi inclusi nella revisione. Le pubblicazioni analizzate differiscono in modo significativo per disegno metodologico, setting assistenziale, specialità clinica, dimensione dei dataset e grado di implementazione delle soluzioni di intelligenza artificiale. Tale variabilità riflette la natura multidimensionale del tema, ma rende complesso il confronto diretto tra i risultati riportati. In pratica, ciascuno studio opera entro un contesto peculiare: ciò che funziona in un reparto di radiologia di un grande ospedale universitario potrebbe non essere replicabile in una piccola struttura sanitaria periferica, e un algoritmo sviluppato su un ampio dataset nazionale può offrire prestazioni molto diverse rispetto a uno addestrato su dati limitati di una singola istituzione. Questa variabilità intrinseca complica qualsiasi tentativo di sintesi generale delle evidenze, poiché manca una base comune e standardizzata su cui valutare in modo omogeneo le diverse soluzioni analizzate.

Analogamente, gli outcome considerati nei vari studi risultano spesso disomogenei e non sempre comparabili tra loro. Molti lavori si concentrano su indicatori di efficienza organizzativa o su misure surrogate (ad esempio la *riduzione dei tempi di attesa*, *l'incremento del tasso di utilizzo delle risorse* o *la diminuzione delle cancellazioni di procedure*), mentre solo una quota più limitata affronta in modo strutturato esiti legati alla qualità clinico-assistenziale, alla sicurezza delle cure o all'impatto sull'esperienza di pazienti e operatori. Questa eterogeneità negli indicatori di risultato rende difficile aggregare o confrontare i dati: un aumento dell'efficienza operativa non equivale necessariamente a un miglioramento clinico se non vengono considerati anche parametri di esito sanitario o di soddisfazione dell'utenza. Ad esempio, la diminuzione dei tempi di attesa per una visita o un intervento è sicuramente un segnale positivo dal punto di vista gestionale, ma non assicura di per sé un migliore esito per il paziente se, parallelamente, non vengono

monitorati aspetti come l'appropriatezza delle cure erogate o la percezione di qualità del servizio.

A fronte di questa disomogeneità di metriche e risultati, risulta complesso trarre conclusioni univoche sull'efficacia complessiva delle soluzioni di IA nei diversi contesti analizzati. Ciascuno studio tende a evidenziare benefici e limiti circoscritti alla propria realtà e agli outcome scelti, impedendo di formulare un giudizio globale che valga trasversalmente in tutti gli ambiti sanitari considerati. In definitiva, l'eterogeneità delle evidenze disponibili impone cautela nell'interpretare i risultati: le indicazioni emerse vanno contestualizzate caso per caso, evitando generalizzazioni che potrebbero non trovare riscontro in situazioni differenti.

6.3 Limiti legati alla qualità dei dati e alle evidenze disponibili

La qualità delle evidenze incluse nella revisione rappresenta un ulteriore elemento critico. Una parte consistente degli studi analizzati si basa infatti su disegni osservazionali, spesso di natura retrospettiva, condotti su database amministrativi o su archivi clinici di singole aziende o istituti. Sebbene tali dataset consentano lo sviluppo e la validazione di modelli predittivi su grandi volumi di informazioni real-world, presentano diversi limiti in termini di affidabilità e rappresentatività. I dati amministrativi e clinici raccolti per scopi operativi possono essere incompleti o imprecisi: ad esempio, non sempre tutte le variabili rilevanti per l'analisi sono disponibili, e possono sussistere errori o omissioni nella registrazione. Inoltre, basare le valutazioni su casistiche provenienti da una singola realtà ospedaliera può introdurre bias di selezione: la popolazione di pazienti e le pratiche assistenziali di quell'istituzione potrebbero non rispecchiare la varietà riscontrabile in altri contesti sanitari. Di conseguenza, un modello di intelligenza artificiale tarato su dati locali potrebbe risultare meno efficace se applicato altrove, dove le condizioni al contorno differiscono. Va anche considerato che l'analisi retrospettiva implica lavorare su dati raccolti in epoche e modalità diverse, spesso senza un disegno preordinato: ciò comporta una certa eterogeneità e variabilità nella qualità dei dati stessi, rendendo più complessa l'interpretazione dei risultati ottenuti.

Un ulteriore aspetto riguarda la carenza di dettagli metodologici riportati in molti studi per quanto concerne la gestione e la preparazione dei dati. Non tutti gli autori

forniscono informazioni esaustive sulle procedure adottate per trattare i dati mancanti o i valori anomali (*outlier*), né sulle tecniche utilizzate per prevenire o correggere eventuali distorsioni nei modelli. Ad esempio, raramente viene specificato se siano state intraprese azioni per mitigare bias algoritmici, come possibili disparità di performance del modello tra diversi sottogruppi di pazienti. L'assenza di tali informazioni rende difficile valutare la solidità dei risultati: senza sapere in che modo il modello sia stato validato e testato in scenari differenti, risulta arduo stimarne la robustezza e l'affidabilità al di fuori del contesto originario. In altre parole, la trasferibilità delle soluzioni di IA proposte rimane incerta. Un algoritmo che mostra elevate prestazioni sui dati di un singolo centro potrebbe non replicare gli stessi risultati in un ambiente diverso, specialmente se vi sono differenze nella popolazione, nelle modalità di raccolta dei dati o nei processi assistenziali. Questa limitata trasparenza sulle modalità di gestione dei dati e sulle strategie di validazione costituisce dunque un limite importante delle evidenze disponibili, poiché impedisce una piena fiducia nell'applicabilità generale dei risultati e ostacola la riproducibilità indipendente degli studi.

6.4 Limitata evidenza sugli esiti clinici e sulla sostenibilità nel tempo

Un limite emerso dalla letteratura analizzata, e di riflesso dalla presente revisione, è la scarsità di evidenze relative agli esiti clinici "hard" e alla sostenibilità delle soluzioni di intelligenza artificiale nel medio-lungo periodo. Gran parte degli studi disponibili si focalizza infatti su risultati organizzativi di breve termine o su simulazioni teoriche, mentre sono molto meno frequenti le analisi che valutino l'impatto diretto dell'IA su esiti clinici concreti. Ad esempio, è comune imbattersi in ricerche che riportano una riduzione dei tempi di attesa o un aumento del volume di pazienti gestiti grazie all'IA, ma è assai raro trovare studi che dimostrino un miglioramento di indicatori clinici quali la durata della degenza ospedaliera, i tassi di complicanze, la mortalità o la qualità di vita dei pazienti. Questa carenza di dati sugli outcome clinici forti implica che attualmente disponiamo di poche prove del fatto che le applicazioni di IA si traducano effettivamente in benefici tangibili per la salute dei pazienti oltre al livello dell'efficienza operativa.

In aggiunta, risultano poco rappresentate in letteratura valutazioni prospettiche estese nel tempo. Pochissimi studi hanno seguito l'implementazione di soluzioni di

IA in contesti reali con disegni prospettici o trial controllati, monitorandone gli effetti nel medio o lungo termine. Ciò significa che non sappiamo se i vantaggi iniziali documentati da alcuni lavori permangano a distanza di anni o se emergano problemi o effetti collaterali con l'uso continuativo della tecnologia. Aspetti fondamentali come l'equità di accesso alle cure e la sicurezza clinica nel tempo sono stati solo marginalmente esplorati: ad esempio, non è quasi mai indagato se l'introduzione di algoritmi di IA possa favorire o penalizzare determinati gruppi di pazienti (*riducendo o ampliando eventuali disparità preesistenti*), né se l'affidamento a sistemi automatizzati di supporto decisionale mantenga nel lungo periodo lo stesso livello di sicurezza e affidabilità senza generare errori imprevisti. L'assenza di queste valutazioni approfondite limita la comprensione dell'effettivo impatto clinico dell'IA e lascia aperti interrogativi sull'eticità e la sicurezza dell'integrazione stabile di tali strumenti nei processi assistenziali.

Analogamente, le analisi economiche e di sostenibilità organizzativa delle soluzioni di IA risultano spesso limitate o del tutto assenti. Molti studi, concentrandosi sulla fase sperimentale o pilota, trascurano di quantificare i costi e le risorse necessarie per portare un sistema basato su IA dalla sperimentazione alla routine clinica quotidiana. Di conseguenza, aspetti come i costi di implementazione (acquisizione di infrastrutture hardware e software, integrazione nei sistemi informativi esistenti), la manutenzione continuativa dei modelli (aggiornamento periodico dei dati, ritaratura degli algoritmi per mantenere le prestazioni), e la formazione del personale sanitario all'uso efficace di questi strumenti rimangono in secondo piano. Senza queste informazioni, risulta difficile valutare il reale rapporto costo-beneficio e il valore aggiunto di una soluzione di IA una volta superata la fase iniziale di adozione. In altre parole, anche laddove vengano dimostrati miglioramenti prestazionali o organizzativi nel breve termine, non è scontato che tali benefici siano sostenibili e giustificati in un'ottica di lungo periodo. La mancanza di evidenze sulla sostenibilità finanziaria e operativa impedisce di comprendere se le innovazioni introdotte possano consolidarsi stabilmente nel sistema sanitario, generando valore duraturo, oppure se rischiano di rimanere interventi isolati e difficilmente scalabili per via di costi elevati o complessità gestionali.

6.5 Rapidità di evoluzione tecnologica e rischio di obsolescenza delle evidenze

Infine, un limite intrinseco a questo ambito di ricerca è rappresentato dalla rapidità con cui evolvono le tecnologie di intelligenza artificiale applicate alla sanità. Il panorama tecnologico è in costante mutamento e progresso: algoritmi, modelli e strumenti che oggi sono considerati innovativi potrebbero diventare obsoleti nel giro di pochi anni (o addirittura mesi) a causa di nuovi sviluppi. Alcune soluzioni descritte negli studi inclusi in questa revisione si collocano a uno stadio di *proof-of-concept* o di sperimentazione iniziale e potrebbero essere rapidamente superate da approcci più avanzati, sia dal punto di vista algoritmico che dell'integrazione nei processi clinici. Ad esempio, un sistema di supporto decisionale sviluppato su tecnologia attuale potrebbe venire soppiantato in breve tempo da nuove piattaforme di intelligenza artificiale capaci di apprendere da quantità di dati molto maggiori o di integrarsi profondamente con i sistemi informativi ospedalieri, rendendo automatica e in tempo reale l'acquisizione dei dati clinici. Questo continuo avanzamento tecnico comporta il rischio che alcune evidenze raccolte divengano presto meno rilevanti o datate: ciò che è stato osservato come efficace o promettente in uno studio condotto qualche anno fa potrebbe non rispecchiare più lo stato dell'arte attuale, semplicemente perché la tecnologia sottostante è evoluta e ha aperto nuove possibilità o nuovi standard di riferimento.

La presente revisione fornisce pertanto una fotografia dello stato dell'arte limitata a un determinato periodo temporale. È importante esserne consapevoli, poiché nuove evidenze e applicazioni potrebbero modificare in tempi relativamente brevi il quadro complessivo emerso dalla presente revisione. In un campo così dinamico, i trend identificati oggi potrebbero essere ribaltati o ulteriormente sviluppati domani: ad esempio, un approccio di IA che attualmente appare marginale potrebbe guadagnare centralità grazie a risultati successivi, oppure, al contrario, una linea di ricerca oggi in voga potrebbe rivelarsi meno promettente del previsto alla luce di nuovi dati. Questo aspetto impone cautela nell'interpretazione dei risultati: le conclusioni tratte devono essere considerate valide alla luce delle conoscenze disponibili fino a questo momento, ma non vanno intese come verità definitive e immutabili. Si sottolinea dunque la necessità di aggiornamenti periodici delle evidenze disponibili: per mantenere una visione accurata e aggiornata dell'impatto

dell'IA in sanità, sarà fondamentale rivedere e integrare regolarmente le conoscenze man mano che il campo progredisce.

In conclusione, la consapevolezza di questi limiti è importante per contestualizzare i risultati della revisione ed evitare interpretazioni eccessivamente ottimistiche o indebite generalizzazioni. Le evidenze raccolte vanno lette in chiave critica, inquadrandole nel contesto specifico e nel periodo temporale in cui sono state prodotte, e tenendo conto delle eterogeneità e delle incertezze metodologiche descritte. Ciò non toglie che il lavoro svolto offra comunque una panoramica utile dell'attuale stato dell'arte e dei principali trend emergenti nell'applicazione dell'IA in ambito sanitario. Le lacune messe in luce indicano al contempo le direttrici verso cui indirizzare gli sforzi futuri: saranno necessari studi più rigorosi e controllati, valutazioni prospettiche sugli esiti clinici e analisi di sostenibilità a lungo termine, oltre a una maggiore attenzione alla qualità dei dati e alla trasparenza metodologica. Solo attraverso un costante approfondimento e un periodico aggiornamento delle conoscenze sarà possibile tradurre pienamente il potenziale dell'intelligenza artificiale in benefici concreti e sostenibili per il sistema sanitario e per i pazienti, superando gli attuali limiti e adeguandosi alle inevitabili evoluzioni tecnologiche.

CAPITOLO 7 - Conclusioni

In qualità di Professionista Sanitario del Servizio Sanitario Nazionale (SSN), ho affrontato questo percorso di tesi non solo come un esercizio accademico, ma come un'occasione concreta di crescita professionale e umana. Il lavoro svolto mi ha portato a riflettere profondamente sul significato della tutela della salute pubblica nel contesto contemporaneo, riaffermando l'importanza dei principi fondanti del nostro sistema sanitario. La Costituzione Italiana, all'articolo 32, afferma che *“La Repubblica tutela la salute come fondamentale diritto dell'individuo e interesse della collettività, e garantisce cure gratuite agli indigenti”*, un principio che ha ispirato la nascita del SSN e che continua a rappresentarne la bussola. Questo richiamo costituzionale ha accompagnato ogni fase della mia ricerca, ricordandomi che dietro dati, tecnologie e innovazioni ci sono sempre le persone, i loro bisogni, e il loro diritto inviolabile a ricevere cure eque, sicure e di qualità.

Nel tracciare le conclusioni di questo lavoro, sento il dovere non solo di valorizzare una visione critica dell'innovazione in sanità, ma anche di affermare con convinzione che l'intelligenza artificiale rappresenta una straordinaria opportunità per il futuro del SSN. L'analisi delle soluzioni di IA in ambito ambulatoriale e chirurgico mi ha mostrato potenzialità incoraggianti nell'ottimizzazione dei percorsi assistenziali, nella gestione delle risorse e nel supporto decisionale clinico. Ma più dei singoli risultati, ciò che emerge con forza è la consapevolezza che tecnologia e umanità non sono dimensioni in opposizione: devono coesistere, rafforzarsi reciprocamente e orientarsi a un fine comune, che è la centralità della persona assistita.

In un sistema come il nostro, fondato su universalità, solidarietà ed equità, l'adozione di strumenti intelligenti non può in alcun modo tradursi in un allontanamento da questi valori. Al contrario, l'IA può e deve diventare un'alleata del SSN, rafforzandone la capacità di rispondere ai bisogni di salute collettivi con maggiore efficienza, tempestività e sostenibilità. Come professionista che opera nell'ambito dei sistemi informativi e della digitalizzazione, credo fermamente che l'IA, se implementata con responsabilità, trasparenza e partecipazione, possa contribuire a rendere la sanità pubblica più resiliente, più equa e più vicina ai bisogni delle persone.

Una delle implicazioni etiche più significative emerse dalla mia riflessione riguarda proprio l'urgenza di preservare la centralità dell'essere umano in ogni processo di innovazione. L'intelligenza artificiale applicata alla sanità solleva questioni complesse: dalla trasparenza degli algoritmi, alla tutela della dignità del paziente, fino alla responsabilità professionale di fronte a sistemi di supporto decisionale automatizzati. Ma è chiaro, oggi più che mai, che l'IA deve rimanere uno strumento nelle mani dei professionisti sanitari e non un sostituto della relazione di cura. Neanche il più sofisticato modello predittivo può replicare, allo stato attuale delle conoscenze, l'empatia, il giudizio etico e la visione globale che il professionista sanitario sviluppa quotidianamente nell'incontro con le persone, senza tuttavia escludere che l'evoluzione tecnologica possa in futuro ridefinire i confini di tali capacità. Tuttavia, proprio per questo, abbiamo l'opportunità di fare dell'IA un'estensione delle nostre competenze, un potenziamento della nostra capacità di prendersi cura, non un ostacolo né una minaccia. La vera sfida non è evitare l'IA, ma **umanizzarla**, costruirla intorno ai bisogni concreti di pazienti e professionisti.

La sicurezza e la qualità delle cure rimangono per me valori non negoziabili. I dati raccolti nella revisione hanno mostrato come le tecnologie intelligenti possano contribuire a ridurre sprechi, prevedere criticità e ottimizzare le agende cliniche; ma questi vantaggi quantitativi vanno sempre letti con lo sguardo rivolto alla qualità dell'assistenza reale. Ottimizzare non deve significare sacrificare la sicurezza. Il ruolo di chi gestisce e dirige le organizzazioni sanitarie è proprio quello di vigilare affinché ogni innovazione venga introdotta nel rispetto degli standard clinici, della sostenibilità organizzativa e del benessere di chi lavora. In questo senso, anche i dati predittivi vanno intesi come supporto alla governance e alla programmazione, non come strumenti di pressione sul paziente o sul professionista.

Parallelamente, la riflessione sull'equità e sulla giustizia sociale ha rappresentato un riferimento etico che ha guidato la mia lettura critica dell'innovazione in sanità. Viviamo in un tempo in cui le disuguaglianze sanitarie rischiano di essere amplificate dalle innovazioni, soprattutto se queste vengono introdotte senza una visione inclusiva. L'IA, come ogni tecnologia, non è neutra: può rafforzare o indebolire i diritti, può colmare o aggravare i divari. Durante questa tesi ho cercato di mantenere

alta l'attenzione sul fatto che l'innovazione non può essere solo accessibile a pochi o concentrata nelle realtà già più avanzate. Serve una progettualità sistemica, che sappia investire anche in alfabetizzazione digitale, infrastrutture e formazione, affinché nessuno venga lasciato indietro. Solo così l'IA potrà essere un vero strumento di equità sociale. Le implicazioni di quanto emerso da questa ricerca si collocano tanto sul piano strategico quanto su quello organizzativo e culturale. Strategicamente, ritengo che l'IA sarà centrale per affrontare le grandi sfide del SSN: dalla carenza di personale all'aumento della domanda di salute. Ma questa transizione richiede una governance attenta, linee guida etiche, investimenti in interoperabilità e una leadership che sappia unire visione, ascolto e capacità di accompagnare il cambiamento. Operativamente, sarà sempre più importante che i professionisti sanitari siano coinvolti nei processi decisionali, formati non solo all'uso delle tecnologie ma anche alla loro interpretazione critica. L'IA, per funzionare bene, ha bisogno di contesti organizzativi maturi, aperti all'innovazione ma ancorati ai principi del miglioramento continuo e della qualità.

In definitiva, credo che il vero significato di questo lavoro risieda nella possibilità di costruire ponti tra due dimensioni che spesso vengono presentate come in tensione: **il progresso tecnologico e la missione umanitaria della cura**. Questa tesi mi ha aiutato a comprendere che non solo possono coesistere, ma devono farlo. L'IA non è il futuro della sanità: è già parte del nostro presente, e siamo noi – professionisti, dirigenti, decisori – a doverne orientare lo sviluppo verso una sanità più intelligente, ma anche più equa, più sicura, più umana.

Concludo con una convinzione personale, innovare in sanità non significa rincorrere la tecnologia, ma orientarla consapevolmente al servizio della persona, della comunità e dei professionisti, preservando i valori che fondano il Servizio sanitario pubblico. Il SSN ha tutte le risorse culturali, etiche e professionali per affrontare questa sfida. Rinnovo quindi il mio impegno, personale e professionale, a contribuire a questo equilibrio virtuoso tra tecnica ed etica, tra dati e diritti, affinché le conquiste della scienza si traducano davvero in **valore per la comunità**, nel rispetto della dignità umana e dei principi costituzionali che continuano a guidare la nostra sanità pubblica.

Bibliografia

- 4ward (2025). *Sanità e intelligenza artificiale, impatti, vantaggi, scenari di una rivoluzione senza precedenti*. <https://www.4ward.it/blog/healthcare-artificial-intelligence-marzo-2025> (ultimo accesso: 01/11/2025)
- Abdelrazig Merghani, A., Ahmed Esmail, A. K., Mubarak Osman, A. M. E., Mohamed, N. A. A., Shentour, S. M. M. A., & Merghani, S. M. A. (2025). *The role of machine learning in management of operating room: A systematic review*. *Cureus*, 17(2), e79400. <https://doi.org/10.7759/cureus.79400>
- AGENAS (2023). *Monitoraggio ex-ante dei tempi di attesa delle prestazioni ambulatoriali – risultati del I semestre 2023*. <https://www.agenas.gov.it/aree-tematiche/comunicazione/primo-piano/2316-evento-agenas-10-novembre-monitoraggio-ex-ante-dei-tempi-di-attesa-delle-prestazioni-ambulatoriali-anno-2023> (ultimo accesso: 01/11/2025)
- AGENAS (2025). *Intelligenza Artificiale – Piattaforma IA per l'assistenza primaria (Missione 6 Salute PNRR)*. <https://www.agenas.gov.it/bandi-di-gara-e-contratti/procedure-di-gara-pnrr/2287-pnrr-m6c1-sub-investimento-1-2-2-4-%E2%80%9Cintelligenza-artificiale%E2%80%9D-procedura-di-dialogo-competitivo-per-la-realizzazione-della-piattaforma-di-intelligenza-artificiale-a-supporto-dell%E2%80%99assistenza-sanitaria-primaria-cig-94572555b6-cup-i81j21000050005> (ultimo accesso: 01/11/2025)
- Al Amin, M., Baldacci, R., & Kayvanfar, V. (2024). *A comprehensive review on operating room scheduling and optimization*. *Operational Research*, 25(3), 1–30. <https://doi.org/10.1007/s12351-024-00884-z>
- Al Zoubi, F., Kashanian, K., Beaulé, P., & Fallavollita, P. (2024). *First deployment of artificial intelligence recommendations in orthopedic surgery*. *Frontiers in Artificial Intelligence*, 7, 1342234. <https://doi.org/10.3389/frai.2024.1342234>
- AlSerkal, Y. M., Ibrahim, N. M., Alsereidi, A. S., Ibrahim, M., Kurakula, S., Naqvi, S. A., Khan, Y., & Oottumadathil, N. P. (2025). *Real-time analytics and AI for managing no-show appointments in primary health care in the United Arab Emirates: Before-and-after study*. *JMIR Formative Research*, 9, e64936. <https://doi.org/10.2196/64936>
- Asouli, J. J., Shao, J., Neifert, S., Gibbs, W. N., Habboub, G., Steinmetz, M. P., Benzell, E., & Mroz, T. E. (2021). *Artificial intelligence and robotics in spine surgery*. *Global Spine Journal*, 11(4), 556–564. <https://doi.org/10.1177/2192568220915718>
- Bar, O., Neimark, D., Zohar, M., Hager, G. D., Girshick, R., Fried, G. M., Wolf, T., & Asselmann, D. (2020). *Impact of data on generalization of AI for surgical intelligence applications*. *Scientific Reports*, 10, 22208. <https://doi.org/10.1038/s41598-020-79173-6>
- Bartek, M. A., Saxena, R. C., Solomon, S., Fong, C. T., Behara, L. D., Venigandla, R., Velagapudi, K., Nair, B. G., & Lang, J. D. (2019). *Improving operating room efficiency: A machine learning approach to predict case-time duration*. *Journal of the American College of Surgeons*, 229(4), 346–354.e3. <https://doi.org/10.1016/j.jamcollsurg.2019.05.029>
- Bellini, V., Guzzon, M., Bigliardi, B., Mordonini, M., Filippelli, S., & Bignami, E. (2020). *Artificial intelligence: A new tool in operating room management. Role of machine learning models in operating room optimization*. *Journal of Medical Systems*, 44(20), 1–10. <https://doi.org/10.1007/s10916-019-1512-1>
- Bellini, V., Russo, M., Domenichetti, T., Panizzi, M., Allai, S., & Bignami, E. G. (2024). *Artificial intelligence in operating room management*. *Journal of Medical Systems*, 48(19). <https://doi.org/10.1007/s10916-024-02038-2>

- Bellini, V., Russo, M., Domenichetti, T., Panizzi, M., Allai, S., & Bignami, E. G. (2024). *Artificial intelligence in operating room management*. *Journal of Medical Systems*, 48(19), 1–16. <https://doi.org/10.1007/s10916-024-02038-2>
- Bignami, E. G., Cozzani, F., Del Rio, P., & Bellini, V. (2021). *The role of artificial intelligence in surgical patient perioperative management*. *Minerva Anestesiologica*, 87(7), 817–822. <https://doi.org/10.23736/S0375-9393.20.14999-X>
- Birkhoff, D. C., van Dalen, A. S. H. M., & Schijven, M. P. (2021). *A review on the current applications of artificial intelligence in the operating room*. *Surgical Innovation*, 28(5), 611–619. <https://doi.org/10.1177/1553350621996961>
- Capaccio, P. (2024). *Dispensa didattica sull'intelligenza artificiale in sanità*. https://www.myecole.it/ooc/wp-content/uploads/2024/06/PW_Capaccio.pdf (ultimo accesso: 05/11/2025)
- Chamarthi, B., Polu, O. R., Anumula, S. K., Ushmani, A., Kasralikar, P., & Syed, A. A. (2025). *Natural language processing (NLP)- and machine learning (ML)-enabled operating room optimization: A PRISMA systematic review anchored in project planning theory*. *Cureus*, 17(4), e82796. <https://doi.org/10.7759/cureus.82796>
- Cittadinanzattiva (2023). *URGENZA SANITÀ: presentato il Rapporto civico sulla salute 2023*. <https://www.cittadinanzattiva.it/comunicati/15741-urgenza-sanita-presentato-da-cittadinanzattiva-il-rapporto-civico-sulla-salute-2023.html> (ultimo accesso: 01/11/2025)
- Coppa, K., Kim, E. J., Oppenheim, M. I., Bock, K. R., Zanos, T. P., & Hirsch, J. S. (2023). *Application of a machine learning algorithm to develop and validate a prediction model for ambulatory non-arrivals*. *Journal of General Internal Medicine*, 38(10), 2298–2307. <https://doi.org/10.1007/s11606-023-08065-y>
- Dashtban, M., & Li, W. (2022). *Predicting non-attendance in hospital outpatient appointments using deep learning approach*. *Health Systems*, 11(3), 189–210. <https://doi.org/10.1080/20476965.2021.1924085>
- Deina, C., Fogliatto, F. S., da Silveira, G. J. C., & Anzanello, M. J. (2024). *Decision analysis framework for predicting no-shows to appointments using machine learning algorithms*. *BMC Health Services Research*, 24(37). <https://doi.org/10.1186/s12913-023-10418-6>
- Entezari, B., Koucheiki, R., Abbas, A., Toor, J., Wolfstadt, J. I., Ravi, B., Whyne, C., & Lex, J. R. (2023). *Improving resource utilization for arthroplasty care by leveraging machine learning and optimization: A systematic review*. *Arthroplasty Today*, 20, 101116. <https://doi.org/10.1016/j.artd.2023.101116>
- Eshghali, M., Kannan, D., Salmanzadeh-Meydani, N., & Esmaieeli Sikaroudi, A. M. (2024). *Machine learning-based integrated scheduling and rescheduling for elective and emergency patients in the operating theatre*. *Annals of Operations Research*, 332, 989–1012. <https://doi.org/10.1007/s10479-023-05168-x>
- Espaillet, A. (2024). *Enhancing operating room surgical efficiency through artificial intelligence: A comprehensive review*. *Surgical Research*, 6(4), 1–8. <https://doi.org/10.33425/2689-1093.1073>
- European Artificial Intelligence Act. (2024). *High-level summary of the AI Act*. <https://artificialintelligenceact.eu/high-level-summary/> (ultimo accesso: 05/11/2025)
- European Commission. (2024). *AI Act | Shaping Europe's digital future*. <https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai> (ultimo accesso: 05/11/2025)
- Fatebenefratelli (2023). *Liste di attesa: i dati sull'andamento dei tempi di attesa*. <https://www.fatebenefratelli.it/blog/liste-attesa> (ultimo accesso: 01/11/2025)
- Feng, H., Jia, Y., Huang, T., Zhou, S., & Chen, H. (2024). *An adaptive decision support system for outpatient appointment scheduling with heterogeneous service times*. *Scientific Reports*, 14(27731). <https://doi.org/10.1038/s41598-024-77873-x>

- Feng, J., Phillips, R. V., Malenica, I., Bishara, A., Hubbard, A. E., Celi, L. A., & Pirracchio, R. (2022). *Clinical artificial intelligence quality improvement: Towards continual monitoring and updating of AI algorithms in healthcare*. *NPJ Digital Medicine*, 5(66). <https://doi.org/10.1038/s41746-022-00611-y>
- Gatto, P. (2025). *Professioni sanitarie e Intelligenza Artificiale: l'Italia fissa i paletti con la Legge 132/2025*. *Dental Tribune Italia*. <https://it.dental-tribune.com/news/professioni-sanitarie-e-intelligenza-artificiale-litalia-fissa-i-paletti-con-la-legge-132-2025/> (ultimo accesso: 05/11/2025)
- Getinge. (2024). *Ora l'IA può aiutarvi a rendere le sale operatorie più produttive*. <https://www.getinge.com/it/approfondimenti/articoli/salute-digitale/intelligenza-artificiale-torin/> (ultimo accesso: 05/11/2025)
- Getinge. (2024). *Utilizzo ottimizzato della sala operatoria con l'intelligenza artificiale*. <https://www.getinge.com/it/approfondimenti/articoli/salute-digitale/utilizzo-ottimizzato-della-sala-operatoria-con-l-intelligenza-artificiale/> (ultimo accesso: 05/11/2025)
- Hofstetter, L., Schweyckart, N., Seiler, C., Brand, C., Rosella, L. C., Farshad, M., Puhan, M. A., & Hincapié, C. A. (2025). *Predicting patient outcomes and risk for revision surgery after hip and knee replacement surgery: Study protocol for a comparison of modelling approaches using the Swiss National Joint Registry (SIRIS)*. *Diagnostic and Prognostic Research*, 9(16). <https://doi.org/10.1186/s41512-025-00200-z>
- Kendale, S., Bishara, A., Burns, M., Solomon, S., Corriere, M., & Mathis, M. (2023). *Machine learning for the prediction of procedural case durations developed using a large multicenter database: Algorithm development and validation study*. *JMIR AI*, 2(1), e44909. <https://doi.org/10.2196/44909>
- Kovoor, J. G., Bacchi, S., Sharma, P., Sharma, S., Kumawat, M., Stretton, B., Gupta, A. K., Chan, W. O., Abou-Hamden, A., & Maddern, G. J. (2024). *Artificial intelligence for surgical services in Australia and New Zealand: Opportunities, challenges and recommendations*. *Medical Journal of Australia*, 220(5), 234–237. <https://doi.org/10.5694/mja2.52225>
- Lee, S.-Y., Eagleson, R. M., Hearld, L. R., Gibson, M. J., Hearld, K. R., Hall, A. G., Burkholder, G. A., McMahon, J., Mahmood, S. Y., Spraberry, C. T., Baker, T. J., Garretson, A. R., Bradley, H. M., & Mugavero, M. J. (2024). *Leveraging machine learning to enhance appointment adherence at a novel post-discharge care transition clinic*. *JAMIA Open*, 7(4), ooae086. <https://doi.org/10.1093/jamiaopen/ooae086>
- Li, L., Diouf, F., & Gorkhali, A. (2022). *Managing outpatient flow via an artificial intelligence-enabled solution*. *Systems Research and Behavioral Science*, 39(3), 415–427. <https://doi.org/10.1002/sres.2870>
- Li, X., Tian, D., Li, W., Dong, B., Wang, H., Yuan, J., Li, B., Shi, L., Lin, X., Zhao, L., & Liu, S. (2021). *Artificial intelligence-assisted reduction in patients' waiting time for outpatient process: A retrospective cohort study*. *BMC Health Services Research*, 21(237). <https://doi.org/10.1186/s12913-021-06248-z>
- Lin, C. K. Y., Ling, T. W. C., & Yeung, W. K. (2017). *Resource allocation and outpatient appointment scheduling using simulation optimization*. *Journal of Healthcare Engineering*, 2017, 9034737. <https://doi.org/10.1155/2017/9034737>
- Loftus, T. J., Tighe, P. J., Filiberto, A. C., Efron, P. A., Brakenridge, S. C., Mohr, A. M., Rashidi, P., Upchurch, G. R., & Bihorac, A. (2020). *Artificial intelligence and surgical decision-making*. *JAMA Surgery*, 155(2), 148–158. <https://doi.org/10.1001/jamasurg.2019.4917>
- Lopes, J., Guimarães, T., Duarte, J., & Santos, M. (2025). *Enhancing surgery scheduling in health care settings with metaheuristic optimization models: Algorithm validation study*. *JMIR Medical Informatics*, 13(1), e57231. <https://doi.org/10.2196/57231>

- Mansoor, M., & Ibrahim, A. F. (2025). *The transformative role of artificial intelligence in plastic and reconstructive surgery: Challenges and opportunities*. *Journal of Clinical Medicine*, 14(8), 2698. <https://doi.org/10.3390/jcm14082698>
- Martinez, O., Martinez, C., Parra, C. A., Rugeles, S., & Suarez, D. R. (2021). *Machine learning for surgical time prediction*. *Computer Methods and Programs in Biomedicine*, 208, 106220. <https://doi.org/10.1016/j.cmpb.2021.106220>
- Marwaha, J. S., Raza, M. M., & Kvedar, J. C. (2023). *The digital transformation of surgery*. *NPJ Digital Medicine*, 6(103). <https://doi.org/10.1038/s41746-023-00846-3>
- MedTech Europe. (2025, 1° agosto). *One year of AI Act: MedTech Europe calls for coherent implementation to unlock the full potential of AI in healthcare* [Comunicato stampa]. <https://www.medtecheurope.org/news-and-events/press-releases/one-year-of-ai-act-medtech-europe-calls-for-coherent-implementation-to-unlock-the-full-potential-of-ai-in-healthcare/> (ultimo accesso: 05/11/2025)
- Mehta, D., Gonzalez, X. T., Huang, G., & Abraham, J. (2024). *Machine learning-augmented interventions in perioperative care: A systematic review and meta-analysis*. *British Journal of Anaesthesia*, 133(6), 1159–1172. <https://doi.org/10.1016/j.bja.2024.08.007>
- Ministero della Salute (2024). *Il decreto sulle liste di attesa è Legge. Schillaci: “Diamo risposte concrete ai cittadini”*. https://www.quotidianosanita.it/governo-e-parlamento/articolo.php?approfondimento_id=19448 (ultimo accesso: 01/11/2025)
- Nasef, D., Nasef, D., Sawiris, V., Weinstein, B., Garcia, J., & Toma, M. (2025). *Integrating artificial intelligence in clinical practice, hospital management, and health policy: Literature review*. *Journal of Hospital Management and Health Policy*, 9(20). <https://doi.org/10.21037/jhmhp-24-138>
- National Institute for Health and Care Excellence. (2024). *Artificial intelligence for patient flow. Horizon scanning report*. <https://www.ncbi.nlm.nih.gov/books/NBK728369/>
- O'Reilly-Shah, V. N., Gentry, K. R., Walters, A. M., Zivot, J., Anderson, C. T., & Tighe, P. J. (2020). *Bias and ethical considerations in machine learning and the automation of perioperative risk assessment*. *British Journal of Anaesthesia*, 125(6), 843–862. <https://doi.org/10.1016/j.bja.2020.07.040>
- Ortiz Luzuriaga, D. R., & Briones Morales, V. E. (2025). *Inteligencia Artificial en la gestión de los servicios de salud: Estado actual y perspectivas futuras*. *ASCE Magazine*, 4(2), 483–502. <https://doi.org/10.70577/ASCE/483.502/2025>
- Ottesen, J. A., Storas, T., Vatnehol, S. A. S., Løvland, G., Vik-Mo, E. O., Schellhorn, T., Skogen, K., Larsson, C., Bjørnerud, A., Groote-Eindbaas, I. R., & Caan, M. W. A. (2025). *Deep learning-based intraoperative MRI reconstruction*. *European Radiology Experimental*, 9(29). <https://doi.org/10.1186/s41747-024-00548-9>
- Pagella Politica (2024). *Vero: 4 milioni di italiani rinunciano a curarsi per le liste d'attesa (e non solo)*. <https://pagellapolitica.it/fact-checking/italiani-rinuncia-cure-liste-attesa> (ultimo accesso: 01/11/2025)
- Parlamento europeo & Consiglio dell'Unione europea. (2024, 13 giugno). *Regolamento (UE) 2024/1689 del Parlamento europeo e del Consiglio del 13 giugno 2024 che stabilisce regole armonizzate sull'intelligenza artificiale e modifica vari regolamenti e direttive (Regolamento sull'intelligenza artificiale – AI Act)*. *Gazzetta ufficiale dell'Unione europea*, L 2024/1689. https://eur-lex.europa.eu/legal-content/IT/TXT/PDF/?uri=OJ:L_202401689 (ultimo accesso: 09/11/2025)
- Regione Emilia-Romagna (2017). *Linee di indirizzo per la gestione delle liste di prenotazione dei ricoveri chirurgici programmati nelle Aziende Sanitarie*. <https://salute.regione.emilia-romagna.it/siseps/liste-di-attesa-ricoveri> (ultimo accesso: 01/11/2025)

- Regione Emilia-Romagna (2024). *Intelligenza artificiale nella sanità del futuro: opportunità e sfide*. <https://salute.regione.emilia-romagna.it/agenda/regione/2024/intelligenza-artificiale-nella-sanita-del-futuro-opportunita-e-sfide> (ultimo accesso: 01/11/2025)
- Rochon, M., Tanner, J., Jurkiewicz, J., Beckhelling, J., Aondoakaa, A., Wilson, K., Dhoonmoon, L., Underwood, M., Mason, L., Harris, R., & Cariaga, K. (2024). *Wound imaging software and digital platform to assist review of surgical wounds using patient smartphones: The development and evaluation of artificial intelligence (WISDOM-AI study)*. *PLOS ONE*, 19(12), e0315384. <https://doi.org/10.1371/journal.pone.0315384>
- Rothenberg, S., Bame, B., & Herskovitz, E. (2022). *Prospective evaluation of a machine-learning prediction model for missed radiology appointments*. *Journal of Digital Imaging*, 35(6), 1690–1693. <https://doi.org/10.1007/s10278-022-00670-3>
- Seo, Y., Jeong, S., Lee, S., Kim, T., Kim, J., Chung, C., Lee, C., Rhee, J. M., Kong, H.-J., & Kim, C.-H. (2024). *Machine-learning-based models for the optimization of post-cervical spinal laminoplasty outpatient follow-up schedules*. *BMC Medical Informatics and Decision Making*, 24(278). <https://doi.org/10.1186/s12911-024-02693-y>
- Shah, V., Yung, H. C., Yang, J., Zaslavsky, J., Algarroba, G. N., Pullano, A., Karpel, H. C., Munoz, N., Aphinyanaphongs, Y., Saraceni, M., Shah, P., Jones, S., & Huang, K. (2024). *Predicting robotic hysterectomy incision time: Optimizing surgical scheduling with machine learning*. *JSL: Journal of the Society of Laparoscopic & Robotic Surgeons*, 28(4), e2024.00040. <https://doi.org/10.4293/JSL.2024.00040>
- Srinivas, S. (2020). *A machine learning-based approach for predicting patient punctuality in ambulatory care centers*. *International Journal of Environmental Research and Public Health*, 17(10), 3703. <https://doi.org/10.3390/ijerph17103703>
- Testa, G. (2025, 14 ottobre). *La legge 23 settembre 2025 n.132 sull'intelligenza artificiale: le disposizioni relative alla sanità. Ius et Salus*, 3/2025 [https://iusetsalus.it/images/inEvidenza/Giacomo TestaLa legge 23 settembre 2025 n 132 sull intelligenza artificiale le disposizioni relative alla sanita.pdf](https://iusetsalus.it/images/inEvidenza/Giacomo%20TestaLa%20legge%2023%20settembre%2025%20n%20132%20sull%20intelligenza%20artificiale%20le%20disposizioni%20relative%20alla%20sanita.pdf) (ultimo accesso: 05/11/2025)
- The Bridge & AGENAS (2023). *Liste attese, dati regionali disomogenei. Al via Osservatorio*. <https://www.sanita33.it/regioni-e-asl/860/liste-attese-dati-regionali-disomogenei-al-via-osservatorio-agenas-the-bridge.html> (ultimo accesso: 01/11/2025)
- Toffaha, K. M., Simsekler, M. C. E., Omar, M. A., & ElKebbi, I. (2025). *Predicting patient no-shows using machine learning: A comprehensive review and future research agenda*. *Intelligence-Based Medicine*, 11, 100229. <https://doi.org/10.1016/j.ibmed.2025.100229>
- Varghese, C., Harrison, E. M., O'Grady, G., & Topol, E. J. (2024). *Artificial intelligence in surgery*. *Nature Medicine*, 30(5), 1257–1268. <https://doi.org/10.1038/s41591-024-02970-3>
- Verma, A., Sanaiha, Y., Hadaya, J., Maltagliati, A. J., Tran, Z., Ramezani, R., Shemin, R. J., & Benharash, P. (2022). *Parsimonious machine learning models to predict resource use in cardiac surgery across a statewide collaborative*. *JTCVS Open*, 11, 214–228. <https://doi.org/10.1016/j.xjon.2022.04.017>
- Wood, K. V., Frings, D., Flood, C., & Thomas, N. (2025). *Artificial intelligence machine learning-driven outpatient appointment management: A qualitative study on acceptability*. *Digital Health*, 11, 1–15. <https://doi.org/10.1177/20552076251321016>
- Woodward, C., Green, J., Reed, M. R., Beard, D. J., & Williams, P. R. (2024). *Risk stratification in hip and knee replacement using Artificial Intelligence: A dual centre study to support the utility of high-volume low-complexity hubs and ambulatory surgery centres*. medRxiv [preprint]. <https://doi.org/10.1101/2024.11.29.24317637>

- World Health Organization. (2021). *Health Ethics & Governance: Big data and artificial intelligence*. <https://www.who.int/teams/health-ethics-governance/emerging-technologies/big-data-and-artificial-intelligence> (ultimo accesso: 05/11/2025)
- World Health Organization. (2021, June 28). *WHO issues first global report on Artificial Intelligence (AI) in health and six guiding principles for its design and use* [Comunicato stampa]. Recuperato da <https://www.who.int/news/item/28-06-2021-who-issues-first-global-report-on-ai-in-health-and-six-guiding-principles-for-its-design-and-use> (ultimo accesso: 05/11/2025)
- World Health Organization. (2023). *Regulatory considerations on artificial intelligence for health*. <https://www.who.int/publications/i/item/9789240078871> (ultimo accesso: 05/11/2025)
- Yousefi, N., Hasankhani, F., Kiani, M., & Yousefi, N. (2020). *Appointment scheduling model in healthcare using clustering algorithms*. *International Journal of Optimization and Control: Theories & Applications*. Retrieved from <https://arxiv.org/abs/1905.03083v3>
- Zain, Z., Almadhoun, M. K. I. K. H., Alsadoun, L., & Bokhari, S. F. H. (2024). *Leveraging artificial intelligence and machine learning to optimize enhanced recovery after surgery (ERAS) protocols*. *Cureus*, 16(3), e56668. <https://doi.org/10.7759/cureus.56668>
- Zaribafzadeh, H., Webster, W. L., Vail, C. J., Daigle, T., Kirk, A. D., Allen, P. J., Henao, R., & Buckland, D. M. (2023). *Development, deployment, and implementation of a machine learning surgical case length prediction model and prospective evaluation*. *Annals of Surgery*, 278(6), 890–895. <https://doi.org/10.1097/SLA.0000000000005936>
- Zhou, Y., Wu, Y., & Zhang, T. (2024). Shaping the future of AI in healthcare through ethics and governance. *Humanities and Social Sciences Communications*, 11(1), Articolo 272. <https://doi.org/10.1057/s41599-024-02894-w>

Appendice A - Dichiarazione di utilizzo di strumenti basati su Intelligenza Artificiale

In un'ottica di trasparenza metodologica e in linea con le recenti raccomandazioni etiche sul corretto impiego dell'intelligenza artificiale nei processi di ricerca e redazione accademica, l'autore dichiara di aver utilizzato alcuni strumenti basati su tecnologie di **Intelligenza Artificiale** (AI) a supporto della presente tesi di laurea.

Nel corso della redazione sono stati impiegati:

- **ChatGPT[®] 5.0, 5.1, 5.2 (OpenAI) e NotebookLM[®] (Google)**, utilizzati per attività di revisione linguistica, riformulazione e sintesi del testo, con l'obiettivo di migliorare chiarezza, coerenza e fluidità espositiva. Tutti i contenuti scientifici, l'impostazione metodologica, le analisi e le discussioni restano integralmente elaborati dall'autore.
- **Elicit[®] e Consensus[®]**, impiegati come strumenti complementari di ricerca bibliografica assistita da AI, a integrazione della metodologia tradizionale di revisione della letteratura (si veda [Capitolo 3.3](#)).
- **Canva[®]**, utilizzato per la realizzazione di grafici e immagini, sia mediante funzionalità tradizionali sia con strumenti basati su AI, al fine di migliorare la qualità visiva del materiale illustrativo.

L'autore conferma di aver mantenuto **piena responsabilità scientifica e critica** in tutte le fasi di produzione, analisi e validazione dei contenuti, verificando l'attendibilità delle informazioni e delle fonti citate.

ALLEGATO A: Griglia di estrazione dati degli studi

ID STUDIO: ST01
TITOLO: Appointment scheduling model in healthcare using clustering algorithms
Titolo in italiano: Modello di pianificazione degli appuntamenti in ambito sanitario mediante algoritmi di clustering

Autore/i	Niloofar Yousefi, Farhad Hasankhani, Mahsa Kiani, Nooshin Yousefi
Anno	2020
Paese	Iran (dataset), USA (affiliazioni autori)
Tipo di studio	Studio osservazionale + sviluppo e valutazione di modello metodologico
Contesto	Ambulatoriale
Specialità/Setting	Prenotazioni ambulatoriali (Shaheed Rajaei Medical and Research Center)
Tecnologia IA	Machine Learning + Ricerca Operativa
Algoritmo/Modello	K-means clustering, Agglomerative hierarchical clustering, DBSCAN, OPTICS; integrazione con Markov Decision Process (MDP) e fluid approximation
Compito/Utilizzo	Classificazione dei pazienti in classi di priorità (clustering) e ottimizzazione dello scheduling degli appuntamenti
Dataset	303 pazienti ambulatoriali real-world, 29 variabili cliniche e anamnestiche – Shaheed Rajaei Medical and Research Center
Outcome principali	Riduzione del tempo di attesa indiretto per pazienti ad alta priorità; miglior allocazione delle agende
Metri di valutazione	Silhouette score per clustering: K-means 0.87, Agglomerative 0.81, DBSCAN 0.92, OPTICS 0.89
Risultati chiave	DBSCAN ha ottenuto la migliore performance di clustering; benchmark policy derivata dall'MDP per allocare appuntamenti entro 7 giorni; allocazione giornaliera differenziata per priorità
Benefici dichiarati	Scheduling più efficiente, riduzione attese per alta priorità, uso ottimale capacità, facilità di implementazione della policy di scheduling
Limiti dichiarati	Applicazione su un singolo dataset di piccole dimensioni; necessità di validazione in altri contesti (studio monocentrico) e su dataset più ampi.
Riferimento Bibliografico	https://arxiv.org/abs/1905.03083v3
Stratificazione per setting	AMBULATORIALE
Inclusione studio nella Revisione	SI

ID STUDIO: ST02
TITOLO: Machine learning for surgical time prediction
Titolo in italiano: Apprendimento automatico per la previsione del tempo chirurgico

Autore/i	Oscar Martinez, Carol Martinez, Carlos A. Parra, Saul Rugeles, Daniel R. Suarez
Anno	2021
Paese	Colombia (Pontificia Universidad Javeriana + Hospital Universitario San Ignacio, Bogotá)
Tipo di studio	Studio osservazionale retrospettivo comparativo su database ospedaliero

Contesto	Chirurgico
Specialità/Setting	25 specialità chirurgiche (con focus su 9 principali: ortopedia e traumatologia, chirurgia generale, ginecologia e ostetricia, neurochirurgia, urologia, chirurgia plastica, otorinolaringoiatria, oftalmologia, chirurgia testa-collo)
Tecnologia IA	Machine Learning (supervised learning)
Algoritmo/Modello	Linear Regression, Support Vector Machines (SVM), Regression Trees, Bagged Trees (ensemble)
Compito/Utilizzo	Previsione della durata dell'intervento chirurgico
Dataset	206.587 record da EHR ospedaliero (2004–2019), puliti a 81.248 record di interventi programmati a procedura singola; fonte real-world
Outcome principali	Accuratezza della predizione dei tempi chirurgici; confronto con metodo basato su esperienza
Metri di valutazione	RMSE (Root Mean Square Error) in minuti; tempi di training e testing
Risultati chiave	Tutti gli algoritmi ML hanno ottenuto un errore di 26–37 min. Bagged Trees: RMSE ≈ 26 min, training 3.16 min, testing 0.49 min (scenario 9 specialità). Bagged Trees ha superato il metodo basato sull'esperienza in 5 su 6 specialità principali, riducendo l'errore fino a 36 min. Peggiori prestazioni per neurochirurgia e plastica (alta variabilità).
Benefici dichiarati	Possibilità di programmare con blocchi più stretti, riduzione tempi morti in sala, miglior uso delle risorse fisiche e aumento qualità del servizio.
Limiti dichiarati	Eliminazione di molteplici procedure e dati incompleti (≈60% record esclusi); peggior performance in specialità con alta variabilità; ML tende a sottostimare i tempi (diversamente dall'approccio esperto che sovrastima, preferito per motivi organizzativi).
Riferimento Bibliografico	https://doi.org/10.1016/j.cmpb.2021.106220
Stratificazione per setting	CHIRURGICO
Inclusione studio nella Revisione	SI

ID STUDIO:	ST03
TITOLO:	Improving Resource Utilization for Arthroplasty Care by Leveraging Machine Learning and Optimization: A Systematic Review
Titolo in italiano:	Migliorare l'utilizzo delle risorse nella cura artroplastica attraverso l'apprendimento automatico e l'ottimizzazione: una revisione sistematica
Autore/i	Bahar Entezari, Robert Koucheiki, Aazad Abbas, Jay Toor, Jesse I. Wolfstadt, Bheeshma Ravi, Cari Whyne, Johnathan R. Lex
Anno	2023
Paese	Canada, con collaborazioni a livello internazionale
Tipo di studio	Revisione sistematica (per la parte predittiva ML) + Revisione narrativa (per la parte di ottimizzazione matematica)
Contesto	Chirurgico
Specialità/Setting	Ortopedia, con focus specifico sull'artroplastica totale di anca e ginocchio (TKA, THA e revisioni TKA)
Tecnologia IA	Machine Learning e Deep Learning, con applicazioni di reti neurali artificiali e utilizzo di diversi algoritmi di apprendimento automatico
Algoritmo/Modello	Gli algoritmi utilizzati comprendono modelli bayesiani, random forest, support vector machine, gradient boosted decision trees, XGBoost, CatBoost, RUSBoost,

Compito/Utilizzo	artificial neural networks, oltre a modelli di regressione lineare, logistica, ridge, lasso ed elastic net L'obiettivo è stato duplice, da un lato la previsione di outcomes clinici quali la durata della degenza, la durata dell'intervento e le riammissioni ospedaliere, dall'altro l'ottimizzazione della programmazione chirurgica tramite strategie matematiche e computazionali
Dataset	I dataset impiegati erano real-world, provenienti sia da singoli centri ospedalieri sia da grandi database nazionali e multinazionali, tra cui SPARCS (New York State), NIS (National Inpatient Sample), OrthoMiDaS OME e NSQIP (American College of Surgeons). Le dimensioni dei campioni variavano da un minimo di 525 a un massimo di 424.443 pazienti
Outcome principali	Gli outcome indagati erano la durata della degenza (LOS), la durata della chirurgia (DOS), le riammissioni entro 90 giorni e l'efficienza dello scheduling chirurgico
Metri di valutazione	Le metriche utilizzate per valutare le performance dei modelli comprendevano AUC con valori compresi tra 0,64 e 0,87 per la previsione di LOS e riammissioni, accuracy che ha raggiunto il 93% nei modelli di regressione logistica e l'85% nelle reti neurali, coefficienti di determinazione R^2 fino a 0,76 per CatBoost rispetto a 0,71 della regressione lineare, e valori di MSE e RMSE con le ANN che hanno ottenuto punteggi migliori (0,69–0,89) rispetto ai modelli di controllo (0,79–0,99) e RMSE intorno a 3,84
Risultati chiave	I modelli di machine learning, in particolare le reti neurali artificiali, hanno generalmente mostrato performance superiori o almeno equivalenti rispetto ai modelli tradizionali di controllo basati su regressioni. Solo in un caso specifico il random forest ha superato la rete neurale corrispondente. Per quanto riguarda l'ottimizzazione, le strategie di programmazione intera mista, le simulazioni Monte Carlo, la discrete-event simulation e i processi decisionali markoviani hanno migliorato significativamente l'efficienza rispetto ai metodi di pianificazione manuale
Benefici dichiarati	Lo studio ha evidenziato come l'uso dell'intelligenza artificiale possa portare a un migliore utilizzo delle risorse ospedaliere, in particolare delle sale operatorie, dei letti e delle recovery room, con una riduzione dei costi, un aumento della produttività e la possibilità di fornire un'assistenza più personalizzata, ad esempio individuando i pazienti a maggiore rischio di riammissione. Inoltre, viene sottolineata la prospettiva di sviluppare modelli di pagamento basati su previsioni personalizzate del consumo di risorse.
Limiti dichiarati	I principali limiti riguardano la scarsa validazione esterna della maggior parte dei modelli, l'assenza di dettagli completi sui metodi di training e validazione, l'elevata eterogeneità delle metriche di outcome utilizzate e la difficoltà di interpretazione dei modelli di deep learning rispetto ai modelli tradizionali. Un ulteriore limite è rappresentato dalla dipendenza da assunzioni di base, come i valori medi di durata degli interventi o della degenza, che possono ridurre l'accuratezza delle ottimizzazioni
Riferimento Bibliografico	https://doi.org/10.1016/j.artd.2023.101116
Stratificazione per setting	CHIRURGICO
Inclusione studio nella Revisione	SI

ID STUDIO:	ST04
TITOLO:	Artificial Intelligence and its role in surgical care in low-income and middle-income countries
Titolo in italiano:	L'intelligenza artificiale e il suo ruolo nell'assistenza chirurgica nei paesi a basso e medio reddito
Autore/i	Ché L. Reddy, Shivani Mitra, John G. Meara, Rifat Atun, Salim Afshar
Anno	2019
Paese	Stati Uniti (Harvard Medical School, Boston), con focus internazionale sui paesi a basso e medio reddito
Tipo di studio	Commento/analisi concettuale (non è una revisione sistematica né uno studio sperimentale)
Contesto	Chirurgico, con analisi a livello di sistemi sanitari nei LMICs
Specialità/Setting	Chirurgia generale e sistemi chirurgici globali, in particolare nei paesi a basso e medio reddito (LMICs)
Tecnologia IA	Intelligenza Artificiale in senso ampio, con menzione di machine learning e deep learning come strumenti applicabili alla gestione dei dati e dei processi sanitari
Algoritmo/Modello	Non vengono descritti algoritmi specifici, ma si discute l'applicazione di approcci di intelligenza artificiale trasparenti e spiegabili a dataset clinici e gestionali
Compito/Utilizzo	Supporto decisionale e ottimizzazione dei sistemi chirurgici, in particolare nell'ambito di governance, workforce, infrastrutture, service delivery e gestione delle informazioni, secondo il quadro NSOAP (National Surgical, Obstetric and Anesthesia Planning)
Dataset	L'articolo non presenta dataset sperimentali; discute l'utilizzo di dati real-world già disponibili nei paesi LMICs (es. Sudafrica, Brasile, India), sottolineando che spesso sono dati asimmetrici, asincroni e frammentati, ma potenzialmente sfruttabili tramite AI
Outcome principali	Non vengono riportati outcome quantitativi; l'articolo sottolinea il potenziale dell'AI nel migliorare efficienza dei processi, distribuzione delle risorse, catene di fornitura, workforce management e capacità di pianificazione chirurgica nei sistemi sanitari dei LMICs
Metri di valutazione	Non applicabili in quanto si tratta di un commento e non di uno studio sperimentale o di validazione algoritmica
Risultati chiave	L'articolo evidenzia come l'AI, sebbene non sia una panacea, possa contribuire a ridurre inefficienze gestionali e organizzative nei sistemi chirurgici dei paesi a basso e medio reddito, supportando la governance, il miglior utilizzo dei dati disponibili e la creazione di sistemi sanitari più resilienti. Sottolinea inoltre la necessità di trasparenza, accountability, etica e regolamentazione nell'adozione di queste tecnologie
Benefici dichiarati	Potenziale miglioramento dell'accesso a cure chirurgiche sicure ed efficaci, ottimizzazione della distribuzione di workforce e infrastrutture, rafforzamento della pianificazione e gestione dei sistemi chirurgici nazionali, riduzione di inefficienze burocratiche e operative
Limiti dichiarati	Mancanza di dati di alta qualità e frammentazione dei dataset nei LMICs, difficoltà legate alla governance, problematiche etiche, questioni di equità e privacy, oltre alla necessità di investimenti in infrastrutture e formazione di competenze locali. Inoltre, non essendo uno studio sperimentale, non fornisce risultati quantitativi o validazioni di modelli concreti
Riferimento Bibliografico	https://doi.org/10.1016/S2589-7500(19)30200-6
Stratificazione per setting	CHIRURGICO
Inclusione studio nella Revisione	NO

ID STUDIO:	ST05
TITOLO:	Parsimonious machine learning models to predict resource use in cardiac surgery across a statewide collaborative
Titolo in italiano:	Modelli parsimoniosi di machine learning per prevedere l'utilizzo delle risorse in cardiocirurgia in un consorzio statale
Autore/i	Arjun Verma, Yas Sanaiha, Joseph Hadaya, Anthony Jason Maltagliati, Zachary Tran, Ramin Ramezani, Richard J. Shemin, Peyman Benharash, per conto dello University of California Cardiac Surgery Consortium
Anno	2022
Paese	Stati Uniti (California, University of California Cardiac Surgery Consortium)
Tipo di studio	Studio osservazionale retrospettivo multicentrico con sviluppo e validazione di modelli predittivi basati su machine learning
Contesto	Chirurgico, cardiocirurgia (CABG e interventi valvolari)
Specialità/Setting	Cardiocirurgia in 5 ospedali accademici della California
Tecnologia IA	Machine Learning con algoritmi basati su alberi decisionali e boosting
Algoritmo/Modello	Gradient Boosted Machines (GBM), Extreme Gradient Boosting (XGBoost), Random Forest, confrontati con modelli di regressione lineare e logistica e con il Society of Thoracic Surgeons (STS) risk score
Compito/Utilizzo	Previsione di outcome clinici e di utilizzo delle risorse, in particolare durata della degenza ospedaliera, durata della degenza in terapia intensiva, mortalità a 30 giorni, insufficienza renale acuta, trasfusioni post-operatorie e reinterventi
Dataset	Dataset real-world derivato dal registro University of California Cardiac Surgery Consortium (UCCSC), con raccolta prospettica di dati clinici. Sono stati inclusi 6.316 pazienti sottoposti a bypass coronarico (CABG) e/o interventi valvolari tra il 2015 e il 2021, di cui 5.028 assegnati al dataset di derivazione e 1.288 a quello di validazione
Outcome principali	Gli outcome primari erano la durata della degenza ospedaliera (LOS) e la durata della degenza in terapia intensiva (ICU LOS). Gli outcome secondari includevano la mortalità a 30 giorni, l'insufficienza renale acuta post-operatoria, la necessità di trasfusioni e i reinterventi
Metri di valutazione	Per le variabili continue è stato utilizzato il coefficiente di determinazione R^2 , con il modello GBM che ha raggiunto valori di 0,42–0,47 per LOS e 0,23 per ICU LOS, superiori rispetto alla regressione lineare. Per le variabili binarie è stata usata l'AUC, con valori pari a 0,69 per la mortalità, 0,76 per l'insufficienza renale acuta, 0,73 per le trasfusioni e 0,80 per i reinterventi. Le performance sono state confrontate anche con lo STS risk score, che ha mostrato AUC più elevate per mortalità e AKI, ma inferiori rispetto ai modelli ML per il reintervento
Risultati chiave	I modelli di machine learning hanno superato i modelli di regressione tradizionali nella previsione della durata della degenza e di diversi esiti clinici post-operatori. Il GBM si è dimostrato il modello complessivamente migliore per LOS e ICU LOS, mentre il Random Forest ha mostrato maggiore capacità discriminativa per le trasfusioni. Sono stati identificati come predittori significativi fattori clinici preoperatori (anemia, disfunzione renale, frazione di eiezione ridotta, comorbidità cardiovascolari), oltre a caratteristiche ospedaliere (volume annuale di interventi, numero di posti letto a bassa intensità, disponibilità di cardiocirurghi)
Benefici dichiarati	Lo studio dimostra che modelli di ML parsimoniosi, basati su un numero limitato di variabili, possono fornire stime accurate e facilmente utilizzabili per guidare la programmazione delle risorse ospedaliere, migliorare la gestione dei posti letto e ottimizzare la pianificazione dei casi chirurgici, soprattutto in situazioni di capacità ridotta come durante la pandemia COVID-19
Limiti dichiarati	La generalizzabilità dei risultati è limitata dal fatto che i dati derivano da soli centri accademici della California, con possibili differenze rispetto ad altre realtà

**Riferimento
Bibliografico
Stratificazione per
setting
Inclusione studio
nella Revisione**

ospedaliera. Inoltre, la mancata inclusione di alcune variabili cliniche e logistiche (ad esempio lo status di trasferimento) potrebbe influenzare le prestazioni predittive. I modelli necessitano di ulteriori validazioni prospettiche ed esterne

<https://doi.org/10.1016/j.xjon.2022.04.017>

SI

CHIRURGICO

ID STUDIO:

ST06

TITOLO:

Predicting patient no-shows using machine learning: A comprehensive review and future research agenda

Titolo in italiano:

Predire le mancate presentazioni dei pazienti utilizzando il machine learning: una revisione completa e un'agenda per le ricerche future

Autore/i

Khaled M. Toffaha, Mecit Can Emre Simsekler, Mohammed Atif Omar, Imad ElKebbi

Anno

2025

Paese

Emirati Arabi Uniti (Khalifa University e altri centri di ricerca in collaborazione)

Tipo di studio

Revisione della letteratura con proposta di agenda di ricerca

Contesto

Ambulatoriale

Specialità/Setting

Servizi sanitari ambulatoriali e gestione degli appuntamenti

Tecnologia IA

Machine Learning, con analisi comparativa di diversi approcci supervisionati

Algoritmo/Modello

Gli studi analizzati hanno impiegato modelli come Random Forest, Support Vector Machine, Gradient Boosted Trees, XGBoost, Regressione Logistica, K-Nearest Neighbors e Reti Neurali Artificiali, evidenziando differenze di performance in base ai dataset

Compito/Utilizzo

Previsione delle mancate presentazioni dei pazienti agli appuntamenti programmati ("no-shows"), con l'obiettivo di ottimizzare l'uso delle risorse e migliorare la gestione dei servizi sanitari

Dataset

La revisione riporta l'uso di dataset real-world provenienti da ospedali e cliniche di vari paesi. Le dimensioni dei dataset sono risultate molto variabili, spaziando da qualche migliaio fino a oltre un milione di appuntamenti schedulati, con dati estratti principalmente da sistemi elettronici di prenotazione e cartelle cliniche ambulatoriali

Outcome principali

L'outcome principale è la capacità predittiva dei modelli nel classificare correttamente i pazienti a rischio di non presentarsi, con l'obiettivo di ridurre i tassi di no-show e i conseguenti sprechi di risorse sanitarie

Metri di valutazione

Le metriche maggiormente riportate negli studi inclusi sono AUC, accuracy, precision, recall e F1-score. In diversi casi i valori di AUC hanno superato 0,80 nei modelli più performanti, in particolare Random Forest e XGBoost, mentre i modelli di regressione logistica hanno generalmente ottenuto performance inferiori

Risultati chiave

I risultati mostrano come i modelli di machine learning abbiano una maggiore capacità predittiva rispetto agli approcci statistici tradizionali, con Random Forest e XGBoost che si distinguono come i più robusti. Le variabili predittive più rilevanti sono risultate i precedenti no-show, le caratteristiche socio-demografiche dei pazienti, la distanza geografica dal centro sanitario, il tempo di attesa tra la prenotazione e l'appuntamento e le modalità di comunicazione con il paziente

Benefici dichiarati	L'uso del machine learning nella previsione dei no-show consente di migliorare la gestione delle agende, ridurre i costi e gli sprechi legati agli appuntamenti mancati, aumentare l'efficienza nell'allocazione del personale e migliorare l'accesso ai servizi sanitari per altri pazienti
Limiti dichiarati	I principali limiti riguardano la mancanza di standardizzazione dei dataset utilizzati, la scarsa disponibilità di dati pubblici e condivisibili, l'eterogeneità dei contesti sanitari e la limitata validazione esterna dei modelli. Vengono inoltre sottolineati rischi etici legati alla privacy e alla gestione dei dati sensibili
Riferimento Bibliografico	https://doi.org/10.1016/j.ibmed.2025.100229
Stratificazione per setting	AMBULATORIALE
Inclusione studio nella Revisione	SI

ID STUDIO:	ST07
TITOLO:	Artificial intelligence machine learning-driven outpatient appointment management: A qualitative study on acceptability
Titolo in italiano:	Gestione degli appuntamenti ambulatoriali guidata dal machine learning: uno studio qualitativo sull'accettabilità

Autore/i	Kerry V. Wood, Daniel Frings, Chris Flood, Nicola Thomas
Anno	2025
Paese	Regno Unito
Tipo di studio	Studio qualitativo basato su interviste
Contesto	Ambulatoriale
Specialità/Setting	Servizi ambulatoriali generali, con focus sulla gestione e ottimizzazione delle agende di appuntamento tramite sistemi di intelligenza artificiale
Tecnologia IA	Machine Learning
Algoritmo/Modello	Lo studio non entra nei dettagli tecnici degli algoritmi, in quanto l'obiettivo non è la validazione tecnica, ma l'esplorazione dell'accettabilità percepita. Viene genericamente descritto un sistema di gestione degli appuntamenti basato su machine learning in grado di prevedere la probabilità di mancata presentazione e di proporre azioni correttive
Compito/Utilizzo	Previsione delle mancate presentazioni ai controlli ambulatoriali e supporto decisionale per ottimizzare la gestione delle agende e ridurre i disservizi organizzativi
Dataset	Non è stato analizzato un dataset quantitativo. La ricerca raccoglie dati qualitativi da pazienti e operatori sanitari coinvolti in servizi ambulatoriali real-world nel Regno Unito
Outcome principali	L'outcome principale era comprendere l'accettabilità percepita da pazienti e professionisti sanitari di un sistema di gestione degli appuntamenti guidato dal machine learning, analizzando i fattori che favoriscono o ostacolano la fiducia, l'adozione e l'integrazione nella pratica clinica
Metri di valutazione	Non applicabili, poiché si tratta di uno studio qualitativo senza analisi predittiva quantitativa Risultati chiave: È emerso che pazienti e operatori riconoscono il potenziale dell'IA nell'aumentare l'efficienza e ridurre i no-show, ma hanno espresso dubbi legati alla trasparenza, alla fiducia verso gli algoritmi e al rischio di riduzione dell'interazione umana. Gli operatori hanno sottolineato la necessità di mantenere il controllo clinico e decisionale, mentre i pazienti hanno evidenziato

Risultati chiave	l'importanza di sentirsi ascoltati e non gestiti esclusivamente da sistemi automatici È emerso che pazienti e operatori riconoscono il potenziale dell'IA nell'aumentare l'efficienza e ridurre i no-show, ma hanno espresso dubbi legati alla trasparenza, alla fiducia verso gli algoritmi e al rischio di riduzione dell'interazione umana. Gli operatori hanno sottolineato la necessità di mantenere il controllo clinico e decisionale, mentre i pazienti hanno evidenziato l'importanza di sentirsi ascoltati e non gestiti esclusivamente da sistemi automatici
Benefici dichiarati	Potenziale riduzione delle mancate presentazioni, miglioramento dell'uso delle risorse, ottimizzazione delle agende e possibilità di un sistema più flessibile e predittivo
Limiti dichiarati	I risultati non sono generalizzabili ad altri contesti poiché derivano da un campione limitato e da un unico setting sanitario nel Regno Unito. Inoltre, non essendo uno studio quantitativo, non fornisce evidenze sulle reali performance predittive dei modelli ML
Riferimento Bibliografico	https://doi.org/10.1177/20552076251321016
Stratificazione per setting	AMBULATORIALE
Inclusione studio nella Revisione	SI

ID STUDIO:	ST08
TITOLO:	Risk stratification in hip and knee replacement using Artificial Intelligence: a dual centre study to support the utility of high-volume low-complexity hubs and ambulatory surgery centres
Titolo in italiano:	Stratificazione del rischio nelle protesi di anca e ginocchio tramite Intelligenza Artificiale: uno studio bicentrico a supporto dell'utilità degli hub ad alto volume e bassa complessità e dei centri chirurgici ambulatoriali

Autore/i	Christopher Woodward, Justin Green, MR Reed, David J. Beard, Paul R. Williams
Anno	2025
Paese	Regno Unito
Tipo di studio	Studio osservazionale retrospettivo e di validazione
Contesto	Chirurgico (protesi ortopediche)
Specialità/Setting	Ortopedia – chirurgia elettiva di anca e ginocchio, con focus su hub ad alto volume/bassa complessità (HVLC) e Ambulatory Surgery Centres
Tecnologia IA	Machine Learning
Algoritmo/Modello	Modello di regressione logistica polinomiale sviluppato e validato su dati clinici
Compito/Utilizzo	Previsione del rischio di complicanze post-operatorie e supporto decisionale per la selezione dei pazienti nei percorsi chirurgici
Dataset	29.658 record di pazienti sottoposti a protesi d'anca e ginocchio in due NHS Trust (Northumbria Healthcare e Swansea Bay University Health Board, UK), periodo 2014–2020. Dataset real-world. Modello sviluppato su 9.137 record e validato su 445 pazienti
Outcome principali	Capacità del modello AI di classificare i pazienti come "alto/moderato rischio" o "basso rischio" per complicanze post-operatorie entro 90 giorni, confrontata con la valutazione clinica tradizionale condotta da anestesisti/consulenti

Metri di valutazione	AUROC, Brier Score, Expected Calibration Error (ECE); Sensibilità 70%, Valore Predittivo Negativo 96%
Risultati chiave	Il modello ha stratificato correttamente il rischio nel 70% dei pazienti con complicanze, con performance comparabile alla valutazione clinica consulente-led. Ha mostrato un'alta capacità di identificare pazienti realmente a basso rischio, con NPV al 96%. Tuttavia, circa il 41% dei pazienti classificati ad alto rischio non ha sviluppato complicanze (falsi positivi)
Benefici dichiarati	Possibile riduzione del carico sulle equipe di valutazione preoperatoria, supporto all'allocazione dei pazienti nei centri HVLC o ACS, miglioramento della pianificazione delle risorse e della sicurezza del paziente, contributo alla gestione delle liste d'attesa post-COVID
Limiti dichiarati	Rischio di falsi positivi elevato; tasso di falsi negativi pari al 30%; risultati validi solo per setting ortopedico e popolazione studiata; preprint non sottoposto a peer review; dipendenza dalla qualità e aggiornamento dei dati clinici
Riferimento Bibliografico	https://doi.org/10.1101/2024.11.29.24317637
Stratificazione per setting	CHIRURGICO
Inclusione studio nella Revisione	SI

ID STUDIO: ST09

TITOLO: **The use of Artificial Intelligence in the out of hospital care settings: A Scoping Review**

Titolo in italiano: **L'uso dell'intelligenza artificiale nei contesti di cura extraospedalieri: una scoping review**

Autore/i	Jamie Miles, Mike Brady, Leanne Smith, Charlotte Cotterill, Charlotte Levey
Anno	2025
Paese	Regno Unito
Tipo di studio	Scoping Review
Contesto	Extraospedaliero / domiciliare / comunitario
Specialità/Setting	Assistenza territoriale e domiciliare, inclusi servizi di community care e long-term care
Tecnologia IA	Machine Learning e Deep Learning, con applicazioni di Natural Language Processing in alcuni ambiti; discussione di approcci ibridi
Algoritmo/Modello	Non viene validato un singolo algoritmo, ma la review raccoglie esperienze da diversi studi che hanno utilizzato Random Forest, Support Vector Machines, Reti Neurali Artificiali, modelli di NLP e sistemi predittivi ibridi
Compito/Utilizzo	Previsione di outcome clinici, ottimizzazione della presa in carico extraospedaliera, supporto decisionale per infermieri e operatori di comunità, monitoraggio a distanza dei pazienti
Dataset	Gli studi inclusi utilizzano dataset eterogenei, per lo più real-world, derivati da cartelle cliniche elettroniche, dispositivi di telemonitoraggio, registri territoriali e dati amministrativi; dimensioni variabili da piccoli studi pilota a dataset su larga scala
Outcome principali	Identificazione delle aree di maggiore impatto dell'IA nell'assistenza extraospedaliera: miglioramento del monitoraggio dei pazienti cronici, supporto alla gestione domiciliare, rilevazione precoce di peggioramenti clinici e ottimizzazione dei percorsi di cura
Metri di valutazione	Variano nei diversi studi inclusi; la review riporta l'uso frequente di AUC,

Risultati chiave	accuracy, precision, recall e F1-score, ma sottolinea la mancanza di uniformità e standardizzazione nelle metriche adottate L'IA si è dimostrata utile nel ridurre i ricoveri ospedalieri, migliorare la continuità assistenziale e supportare gli operatori sanitari nella gestione di pazienti complessi al domicilio. Tuttavia, la qualità degli studi varia notevolmente e la maggior parte delle applicazioni si trova ancora in fase sperimentale o pilota
Benefici dichiarati	Maggiore efficienza nell'uso delle risorse sanitarie, migliore continuità assistenziale, supporto alle decisioni cliniche in contesti extraospedalieri, possibilità di personalizzare la presa in carico e migliorare l'aderenza terapeutica dei pazienti cronici
Limiti dichiarati	Scarsità di studi clinici di alta qualità e di validazioni prospettiche; forte eterogeneità metodologica; limitata trasferibilità dei risultati; problematiche etiche e legate alla privacy dei dati; necessità di formare adeguatamente il personale per integrare questi strumenti nella pratica quotidiana
Riferimento Bibliografico	https://doi.org/10.1101/2025.04.04.25325245
Stratificazione per setting	EXTRA-OSPEDALIERO
Inclusione studio nella Revisione	NO

ID STUDIO:	ST10
TITOLO:	Integrating artificial intelligence in clinical practice, hospital management, and health policy: literature review
Titolo in italiano:	Integrazione dell'intelligenza artificiale nella pratica clinica, nella gestione ospedaliera e nelle politiche sanitarie: revisione della letteratura
Autore/i	Daniel Nasef, Demarcus Nasef, Viola Sawiris, Brett Weinstein, Jodan Garcia, Milan Toma
Anno	2025
Paese	Stati Uniti (con collaborazioni accademiche internazionali)
Tipo di studio	Revisione Narrativa della letteratura
Contesto	Entrambi, con particolare attenzione a scenari trasversali clinici, gestionali e di policy
Specialità/Setting	Applicazioni generali di intelligenza artificiale in sanità (diagnostica, clinica, management ospedaliero, telemedicina, analisi dei dati per le politiche sanitarie)
Tecnologia IA	Machine Learning, Deep Learning, Natural Language Processing, approcci ibridi
Algoritmo/Modello	La review non si concentra su un singolo modello, ma raccoglie applicazioni che includono reti neurali profonde, algoritmi di classificazione, modelli NLP per estrazione di informazioni cliniche, sistemi predittivi ibridi per la gestione ospedaliera
Compito/Utilizzo	Supporto decisionale clinico, previsione di outcome, ottimizzazione dei processi ospedalieri, supporto alla definizione di policy sanitarie
Dataset	Gli studi inclusi utilizzano dataset variabili, prevalentemente real-world (EHR, registri sanitari, dati amministrativi e clinici); la review sottolinea l'eterogeneità delle fonti e la difficoltà di standardizzazione
Outcome principali	Identificazione dei benefici dell'IA in tre aree chiave: miglioramento delle decisioni cliniche, efficientamento della gestione ospedaliera, e impatto sulla definizione di politiche sanitarie basate sui dati
Metri di valutazione	Variano in base agli studi analizzati; frequentemente riportati AUC, accuracy, F1-

Risultati chiave	score, ma non uniformi tra i lavori 'IA ha dimostrato di migliorare l'accuratezza diagnostica, ottimizzare l'allocatione delle risorse ospedaliere e fornire evidenze utili alle politiche sanitarie. Tuttavia, permangono sfide legate alla mancanza di trasparenza algoritmica, ai rischi etici e alla difficoltà di integrare i modelli nella pratica quotidiana
Benefici dichiarati	Aumento dell'efficienza clinica, riduzione dei costi gestionali, possibilità di decisioni politiche più informate e data-driven, maggiore personalizzazione delle cure
Limiti dichiarati	Mancanza di standardizzazione dei dataset, rischi di bias algoritmico, limitata applicabilità clinica reale di alcuni modelli, scarsa validazione prospettica, problematiche etiche e di privacy
Riferimento Bibliografico	https://jhmhp.amegroups.org/article/view/9575/html
Stratificazione per setting	SI
Inclusione studio nella Revisione	AMBITO GENERALISTA AI IN SANITA'

ID STUDIO:	ST11
TITOLO:	Prospective Evaluation of a Machine-Learning Prediction Model for Missed Radiology Appointments
Titolo in italiano:	Valutazione prospettica di un modello predittivo di machine learning per appuntamenti radiologici mancati

Autore/i	Steven Rothenberg, Bill Bame, Ed Herskovitz
Anno	2022
Paese	Stati Uniti
Tipo di studio	Studio osservazionale prospettico con validazione di modello predittivo
Contesto	Ambulatoriale (radiologia)
Specialità/Setting	Radiologia diagnostica in un centro ambulatoriale ospedaliero
Tecnologia IA	Machine Learning
Algoritmo/Modello	Gradient Boosted Regression Trees (XGBoost, CatBoost testato ma non adottato). Confronto retrospettivo anche con Logistic Regression, Ridge Regression, Random Forest, Deep Feedforward Neural Network, Support Vector Regression
Compito/Utilizzo	Previsione del rischio di mancata presentazione (no-show) agli appuntamenti radiologici
Dataset	4.546.104 appuntamenti programmati (631.386 no-show) dal sistema sanitario University of Maryland (2017–2020). Validazione prospettica su 2.264 appuntamenti in un singolo centro ambulatoriale di radiologia. Dataset real-world derivato da EMR
Outcome principali	Capacità del modello ML di identificare pazienti ad alto rischio di no-show rispetto alla media, con valutazione sia retrospettiva sia prospettica
Metri di valutazione	AUC retrospettivo 0.93; AUC prospettico 0.73 (p < 0.0005). Analisi per sottogruppi con rischio relativo fino a 6.08 nei pazienti ad alto rischio (>0.30) rispetto a quelli a basso rischio (<0.10)
Risultati chiave	Il modello ha mostrato ottima performance retrospettiva (AUC 0.93), ma prestazioni inferiori nella validazione prospettica (AUC 0.73). I pazienti ad alto rischio avevano una probabilità tripla di no-show rispetto agli altri. Alcuni predittori forti: numero precedente di no-show, flag "interpreter needed", giorno della settimana, dipartimento, lead time dell'appuntamento

Benefici dichiarati	Possibilità di identificare in anticipo i pazienti a rischio per programmare interventi mirati (es. reminder, trasporto), migliorando l'uso delle risorse e riducendo i disservizi organizzativi
Limiti dichiarati	Performance peggiore in prospettiva rispetto a retrospettiva; rischio di bias legato a fattori socio-demografici (razza, codice postale, genere); necessità di aggiornare il modello con dati post-COVID per maggiore robustezza; ROI limitato per interventi costosi, dato il basso tasso assoluto di no-show
Riferimento Bibliografico	https://doi.org/10.1007/s10278-022-00670-3
Stratificazione per setting	AMBULATORIALE RADIOLOGIA
Inclusione studio nella Revisione	SI

ID STUDIO: **ST12**

TITOLO: **Artificial Intelligence in Operating Room Management**

Titolo in italiano: **Intelligenza artificiale nella gestione della sala operatoria**

Autore/i	Valentina Bellini, Michele Russo, Tania Domenichetti, Matteo Panizzi, Simone Allai, Elena Giovanna Bignami
Anno	2024
Paese	Italia
Tipo di studio	Revisione sistematica della letteratura
Contesto	Chirurgico
Specialità/Setting	Anestesia e gestione perioperatoria, con focus sulla pianificazione e l'organizzazione delle sale operatorie
Tecnologia IA	Machine Learning, Deep Learning, Reinforcement Learning, Natural Language Processing
Algoritmo/Modello	Discussione di diversi approcci riportati in letteratura: Random Forest, Support Vector Machines, Convolutional Neural Networks, modelli di Reinforcement Learning per l'ottimizzazione delle agende operatorie e NLP per analisi di testi clinici
Compito/Utilizzo	Ottimizzazione della programmazione operatoria, previsione di complicanze, supporto decisionale in anestesia e perioperatorio, gestione delle risorse umane e materiali in sala operatoria
Dataset	Gli studi analizzati riportano dataset real-world derivati da registri chirurgici ospedalieri, cartelle cliniche elettroniche e sistemi di prenotazione operatoria. Dimensione variabile da piccoli campioni locali a grandi dataset multicentrici; in alcuni casi impiego di dati simulati per scenari di ottimizzazione
Outcome principali	Identificazione delle principali aree in cui l'IA può migliorare la gestione del blocco operatorio: scheduling, predizione dei tempi chirurgici, riduzione delle cancellazioni e ottimizzazione delle risorse. Evidenziati anche potenziali benefici in termini di sicurezza del paziente e qualità dell'assistenza perioperatoria
Metri di valutazione	Variano nei lavori inclusi: AUC, accuracy, MAE e RMSE sono le più riportate; miglioramenti significativi nella predizione dei tempi operatori e nella classificazione del rischio perioperatorio rispetto ai modelli tradizionali
Risultati chiave	L'IA applicata alla sala operatoria si è dimostrata promettente per aumentare l'efficienza organizzativa, ridurre sprechi e tempi morti, migliorare la previsione dei tempi chirurgici e la sicurezza perioperatoria. Tuttavia, le evidenze sono ancora frammentarie e spesso limitate a studi preliminari o single-center
Benefici dichiarati	Miglior utilizzo delle risorse, riduzione delle liste d'attesa, maggiore accuratezza

Limiti dichiarati	nella programmazione, miglioramento della qualità assistenziale e riduzione del carico di lavoro per il personale sanitario Mancanza di standardizzazione dei dataset, variabilità dei modelli, limitata trasferibilità dei risultati in contesti diversi, scarsa validazione prospettica e necessità di studi multicentrici di più ampie dimensioni
Riferimento Bibliografico	https://doi.org/10.1007/s10916-024-02038-2
Stratificazione per setting	CHIRURGICO
Inclusione studio nella Revisione	SI

ID STUDIO:	ST13
TITOLO:	Application of a Machine Learning Algorithm to Develop and Validate a Prediction Model for Ambulatory Non-Arrivals
Titolo in italiano:	Applicazione di un algoritmo di machine learning per sviluppare e validare un modello predittivo delle mancate presentazioni ambulatoriali

Autore/i	Kevin Coppa, Eun Ji Kim, Michael I. Oppenheim, Kevin R. Bock, Theodoros P. Zanos, Jamie S. Hirsch
Anno	2023
Paese	Stati Uniti
Tipo di studio	Studio osservazionale retrospettivo con sviluppo e validazione di modello predittivo
Contesto	Ambulatoriale
Specialità/Setting	Appuntamenti ambulatoriali multipli in un grande sistema ospedaliero statunitense
Tecnologia IA	Machine Learning
Algoritmo/Modello	Gradient Boosting Machine (GBM), confrontato con Regressione Logistica
Compito/Utilizzo	Previsione delle mancate presentazioni (non-arrivals) agli appuntamenti ambulatoriali
Dataset	1.732.562 appuntamenti consecutivi programmati in un ospedale accademico di New York (2017–2019), dataset real-world. Suddiviso in training (60%), validation (20%) e test (20%)
Outcome principali	Capacità del modello ML di identificare i pazienti a rischio di non presentarsi agli appuntamenti, rispetto al modello di regressione logistica tradizionale
Metri di valutazione	AUC ROC: 0.768 per GBM vs 0.739 per regressione logistica ($p < 0.001$); precision-recall curve superiore per GBM; net reclassification improvement statistic significativo
Risultati chiave	Il modello ML ha superato la regressione logistica nella predizione delle mancate presentazioni. Fattori predittivi più rilevanti: numero precedente di non-arrivals, distanza dall'ospedale, tempo di attesa dell'appuntamento, tipo di visita e status assicurativo. Lo studio dimostra che ML può essere implementato per ottimizzare interventi mirati
Benefici dichiarati	Migliore accuratezza predittiva rispetto ai modelli tradizionali; possibilità di allocare interventi (promemoria, supporto logistico) ai pazienti a rischio; incremento dell'efficienza organizzativa e riduzione dei costi legati alle mancate presentazioni
Limiti dichiarati	Studio single-center con dati di un unico sistema ospedaliero; rischio di bias legati a variabili socioeconomiche; assenza di validazione esterna; possibilità di

**Riferimento
Bibliografico
Stratificazione per
setting
Inclusione studio
nella Revisione**

overfitting se non aggiornato regolarmente

<https://doi.org/10.1007/s11606-023-08065-y>

AMBULATORIALE

SI

ID STUDIO:

ST14

TITOLO:

Machine-learning-based models for the optimization of post-cervical spinal laminoplasty outpatient follow-up schedules

Titolo in italiano:

Modelli basati su machine learning per l'ottimizzazione dei follow-up ambulatoriali post-laminoplastica cervicale

Autore/i

Yechan Seo, Seoi Jeong, Siyoung Lee, Tae-Shin Kim, Jun-Hoe Kim, Chun Kee Chung, Chang-Hyun Lee, John M. Rhee, Hyoun-Joong Kong, Chi Heon Kim

Anno

2024

Paese

Corea del Sud (con collaborazioni internazionali, inclusi Stati Uniti)

Tipo di studio

Studio osservazionale retrospettivo con sviluppo e validazione di modelli predittivi

Contesto

Chirurgico (con gestione ambulatoriale post-operatoria)

Specialità/Setting

Neurochirurgia spinale – follow-up post-intervento di laminoplastica cervicale

Tecnologia IA

Machine Learning

Algoritmo/Modello

Gradient Boosting Machine (XGBoost), Random Forest, Support Vector Machine, Reti Neurali Artificiali; confrontati con Regressione Logistica

Compito/Utilizzo

Previsione dei pazienti che necessitano di follow-up più intensivi; ottimizzazione della frequenza e della programmazione degli appuntamenti ambulatoriali post-operatori

Dataset

682 pazienti sottoposti a laminoplastica cervicale tra il 2009 e il 2019 presso il Seoul National University Hospital. Dataset real-world retrospettivo, con suddivisione in training e test set

Outcome principali

Capacità dei modelli ML di identificare i pazienti a rischio di complicanze tardive o di necessità di follow-up più ravvicinati, rispetto alla programmazione standard

Metri di valutazione

AUC ROC fino a 0.84 per XGBoost (modello migliore), superiori ai modelli tradizionali. Reportate anche accuracy, precision e recall, con valori migliori per Random Forest e XGBoost

Risultati chiave

I modelli ML, in particolare XGBoost, hanno migliorato la predizione dei pazienti che richiedevano follow-up più intensivi, consentendo una riduzione dei controlli non necessari senza compromettere la sicurezza clinica. Le variabili predittive più importanti erano punteggi clinici post-operatori, dati demografici e caratteristiche chirurgiche

Benefici dichiarati

Possibilità di personalizzare i follow-up post-operatori, riducendo il carico su pazienti e ospedali, ottimizzando l'uso delle risorse ambulatoriali e migliorando l'efficienza complessiva dei percorsi post-chirurgici

Limiti dichiarati

Studio monocentrico, con popolazione omogenea (coreana) e potenziale riduzione della generalizzabilità. Analisi retrospettiva senza validazione prospettica esterna. Necessità di ulteriori studi multicentrici per confermare l'applicabilità

**Riferimento
Bibliografico**

<https://doi.org/10.1186/s12911-024-02693-y>

Stratificazione per setting	SI
Inclusione studio nella Revisione	AMBULATORIALE
ID STUDIO:	ST15
TITOLO:	Artificial intelligence-assisted reduction in patients' waiting time for outpatient process: a retrospective cohort study
Titolo in italiano:	Riduzione dei tempi di attesa dei pazienti nei percorsi ambulatoriali tramite intelligenza artificiale: uno studio di coorte retrospettivo
Autore/i	Xiaoqing Li, Dan Tian, Weihua Li, Bin Dong, Hansong Wang, Jiajun Yuan, Biru Li, Lei Shi, Xulin Lin, Liebin Zhao, Shijian Liu
Anno	2021
Paese	Cina
Tipo di studio	Studio osservazionale retrospettivo di coorte
Contesto	Ambulatoriale
Specialità/Setting	Gestione dei percorsi ambulatoriali ospedalieri (processo di prenotazione e attesa)
Tecnologia IA	Machine Learning e ottimizzazione algoritmica basata su AI
Algoritmo/Modello	Algoritmo predittivo-ottimizzativo sviluppato ad hoc per la gestione dei flussi ambulatoriali; non viene riportato un singolo modello ML ma un sistema AI di scheduling intelligente
Compito/Utilizzo	Ottimizzazione dei flussi ambulatoriali e riduzione dei tempi di attesa dei pazienti
Dataset	Dati real-world provenienti dal sistema informativo ospedaliero; 75.284 pazienti ambulatoriali analizzati in un grande ospedale cinese tra 2018 e 2019
Outcome principali	Confronto tra percorso tradizionale e percorso AI-assistito in termini di tempi di attesa complessivi per le visite ambulatoriali
Metri di valutazione	Riduzione media dei tempi di attesa pari al 16,9% (da 2,18 ore a 1,81 ore, $p < 0.001$). Miglioramento significativo soprattutto nelle fasce orarie di picco
Risultati chiave	L'introduzione dell'AI ha permesso di ridurre in modo statisticamente significativo i tempi di attesa per i pazienti ambulatoriali, migliorando il flusso dei percorsi clinici e aumentando la soddisfazione dei pazienti. L'efficienza organizzativa è stata particolarmente evidente nei reparti con maggior afflusso
Benefici dichiarati	Miglioramento dell'efficienza del sistema ambulatoriale, riduzione dei tempi di attesa, maggiore soddisfazione dei pazienti e ottimizzazione delle risorse ospedaliere
Limiti dichiarati	Studio monocentrico, retrospettivo, con potenziale riduzione della generalizzabilità dei risultati. Non è stato possibile isolare tutti i fattori esterni che influenzano i tempi di attesa (es. stagionalità, personale disponibile). Mancanza di un trial prospettico randomizzato
Riferimento Bibliografico	https://doi.org/10.1186/s12913-021-06248-z
Stratificazione per setting	SI
Inclusione studio nella Revisione	AMBULATORIALE

ID STUDIO:	ST16
TITOLO:	Development, deployment and scaling of operating room-ready artificial intelligence for real-time surgical decision support
Titolo in italiano:	Sviluppo, implementazione e scalabilità di un'intelligenza artificiale pronta per la sala operatoria per il supporto decisionale chirurgico in tempo reale
Autore/i	Sergey Protserov, Jaryd Hunter, Haochi Zhang, Pouria Mashouri, Caterina Masino, Michael Brudno, Amin Madani
Anno	2024
Paese	Canada (University of Toronto, Sinai Health, Vector Institute)
Tipo di studio	Studio di sviluppo e implementazione con validazione in ambiente reale
Contesto	Chirurgico
Specialità/Setting	Chirurgia generale e addominale, con test in sala operatoria reale
Tecnologia IA	Deep Learning (Computer Vision)
Algoritmo/Modello	Reti Neurali Convolutionali (CNN) ottimizzate per analisi video in tempo reale durante procedure chirurgiche; pipeline AI containerizzata per deployment su infrastruttura ospedaliera
Compito/Utilizzo	Supporto decisionale in tempo reale durante la chirurgia laparoscopica, riconoscimento di strutture anatomiche critiche, avvisi di sicurezza intraoperatori
Dataset	Oltre 300 ore di video chirurgici laparoscopici annotati manualmente (dataset real-world multicentrico). Validazione e deployment effettuati in più ospedali del network di Toronto
Outcome principali	Dimostrazione della fattibilità tecnica di deployment e utilizzo di modelli AI in sala operatoria in tempo reale; valutazione delle prestazioni in termini di accuratezza, latenza e robustezza del sistema
Metri di valutazione	Accuracy del riconoscimento anatomico superiore all'85%, latenza inferiore a 150 ms, robustezza testata su diversi ambienti chirurgici. Reportati anche F1-score e precision-recall in range 0.80–0.90
Risultati chiave	L'AI è stata implementata con successo in sala operatoria senza interrompere i flussi chirurgici; il sistema ha fornito feedback in tempo reale con accuratezza elevata, dimostrando potenziale per migliorare sicurezza e decision-making chirurgico. È stato validato il modello di scalabilità attraverso più ospedali
Benefici dichiarati	Miglioramento della sicurezza chirurgica, riduzione degli errori intraoperatori, possibilità di standardizzare la qualità delle cure chirurgiche, trasferibilità della tecnologia in diversi ospedali
Limiti dichiarati	Necessità di ulteriori validazioni multicentriche su larga scala; rischio di bias legati ai dataset annotati; sfide infrastrutturali e di integrazione con i sistemi ospedalieri; questioni etiche e legali relative all'uso di AI in chirurgia
Riferimento Bibliografico	https://doi.org/10.1038/s41746-024-01225-2
Stratificazione per setting	CHIRURGICO
Inclusione studio nella Revisione	NO

ID STUDIO: ST17
TITOLO: Deep learning-based Intraoperative MRI reconstruction
Titolo in italiano: Ricostruzione intraoperatoria della risonanza magnetica basata su deep learning

Autore/i	Jon André Ottesen, Tryggve Storas, Svein Are Sirirud Vatnehol, Grethe Løvland, Einar Osland Vik-Mo, Till Schellhorn, Karoline Skogen, Christopher Larsson, Atle Bjørnerud, Inge Rasmus Groote-Eindbaas, Matthan W. A. Caan
Anno	2025
Paese	Norvegia (con collaborazioni internazionali nei Paesi Bassi e Svezia)
Tipo di studio	Studio osservazionale retrospettivo con confronto diagnostico
Contesto	Chirurgico (neurochirurgia intraoperatoria)
Specialità/Setting	Neuroradiologia e neurochirurgia, imaging intraoperatorio in pazienti con tumori cerebrali
Tecnologia IA	Deep Learning
Algoritmo/Modello	Densely Interconnected Residual Cascading Network (DIRCN), addestrato sul dataset fastMRI neuro
Compito/Utilizzo	Miglioramento della qualità delle immagini MRI intraoperatorie, supporto decisionale per neurochirurghi durante la resezione tumorale
Dataset	40 pazienti sottoposti a resezione di tumore cerebrale (glioblastomi, astrocitomi, oligodendrogliomi, metastasi, linfomi ecc.) tra novembre 2021 e giugno 2023; dataset real-world intraoperatorio
Outcome principali	Confronto tra ricostruzione con deep learning (DL) e metodo tradizionale compressed sensing (CS) in termini di qualità, contrasto, artefatti e utilità clinica intraoperatoria
Metri di valutazione	Valutazioni qualitative su scala Likert 1–5 da due neuroradiologi e un neurochirurgo; analisi percentuale di preferenza DL vs CS
Risultati chiave	I neuroradiologi hanno preferito le immagini DL in 33/40 e 39/40 casi, mentre il neurochirurgo ha preferito CS in 18/40 casi e DL in 8/40. DL ha migliorato risoluzione e SNR, ma ha introdotto artefatti di striping e riduzione del segnale lontano dalle bobine
Benefici dichiarati	Maggiore qualità diagnostica percepita dai neuroradiologi, potenziale riduzione dei tempi di acquisizione, miglior supporto alle decisioni intraoperatorie
Limiti dichiarati	Presenza di artefatti specifici del DL; differenze di percezione tra radiologi e chirurghi; studio monocentrico e retrospettivo con campione ridotto; necessità di ulteriori validazioni prospettiche multicentriche
Riferimento Bibliografico	https://doi.org/10.1186/s41747-024-00548-9
Stratificazione per setting	CHIRURGICO
Inclusione studio nella Revisione	SI

ID STUDIO: ST18
TITOLO: Machine learning in perioperative medicine: a systematic review
Titolo in italiano: Machine learning in medicina perioperatoria: una revisione sistematica

Autore/i Valentina Bellini, Marina Valente, Giorgia Bertorelli, Barbara Pifferi, Michelangelo Craca, Monica Mordonini, Gianfranco Lombardo, Eleonora Bottani, Paolo Del Rio, Elena Bignami

Anno	2022
Paese	Italia
Tipo di studio	Revisione sistematica
Contesto	Chirurgico (medicina perioperatoria)
Specialità/Setting	Anestesia, chirurgia generale, gestione perioperatoria
Tecnologia IA	Machine Learning, con applicazioni anche di Deep Learning
Algoritmo/Modello	Diversi modelli ML analizzati nei 53 studi inclusi: Random Forest, Support Vector Machine, Gradient Boosting, Reti Neurali Convolutionali e Ricorrenti, modelli di regressione multivariata
Compito/Utilizzo	Previsione di complicanze perioperatorie, outcome chirurgici, gestione anestesiológica e ottimizzazione delle risorse perioperatorie
Dataset	Variabili per ciascuno degli studi inclusi; dataset real-world principalmente derivati da registri chirurgici, cartelle cliniche elettroniche e database multicentrici
Outcome principali	Identificazione del potenziale del ML nel migliorare la predizione di complicanze e outcome perioperatori, con performance generalmente superiori ai modelli statistici tradizionali
Metri di valutazione	Le metriche più frequentemente riportate sono AUC, accuracy, precision, recall e RMSE. Molti studi hanno mostrato miglioramenti significativi delle AUC (fino a >0.85) rispetto ai modelli standard
Risultati chiave	Il ML ha dimostrato di poter predire meglio degli score clinici tradizionali complicanze come insufficienza respiratoria, infezioni post-operatorie, necessità di riammissione e mortalità. Tuttavia, i risultati sono eterogenei e non sempre trasferibili nella pratica clinica
Benefici dichiarati	Possibilità di migliorare la sicurezza perioperatoria, ottimizzare la gestione delle risorse, supportare la personalizzazione della cura, ridurre costi e inefficienze
Limiti dichiarati	Alta eterogeneità metodologica degli studi inclusi; mancanza di validazione prospettica; rischio di overfitting; scarsa interpretabilità clinica di alcuni modelli ML
Riferimento Bibliografico	https://doi.org/10.1186/s44158-022-00033-y
Stratificazione per setting	CHIRURGICO
Inclusione studio nella Revisione	SI

ID STUDIO:	ST19
TITOLO:	Leveraging Artificial Intelligence and Machine Learning to Optimize Enhanced Recovery After Surgery (ERAS) Protocols
Titolo in italiano:	Sfruttare l'intelligenza artificiale e il machine learning per ottimizzare i protocolli ERAS (Enhanced Recovery After Surgery)
Autore/i	Zukhruf Zain; Mohammed Khaleel I. K.H. Almadhoun; Lara Alsadoun; Syed Faqeer Hussain Bokhari. Gli autori afferiscono a Aga Khan University Hospital (Karachi, Pakistan), Mutah University (Karak, Giordania), Chelsea and Westminster Hospital (Londra, Regno Unito) e King Edward Medical University (Lahore, Pakistan)
Anno	2024
Paese	Multicentrico internazionale (Pakistan, Giordania, Regno Unito), in base alle affiliazioni riportate.
Tipo di studio	Editoriale/Commento (Cureus – sezione "Editorial").

Contesto	Chirurgico, con focus sulla medicina perioperatoria e sull'implementazione dei protocolli ERAS.
Specialità/Setting	Assistenza perioperatoria trasversale alle specialità chirurgiche (pre-, intra- e post-operatorio) secondo i pilastri ERAS (ottimizzazione preoperatoria, chirurgia miniminvasiva, cura postoperatoria standardizzata).
Tecnologia IA	Intelligenza artificiale e machine learning in senso ampio, con enfasi su modelli predittivi e strumenti per personalizzazione degli interventi; vengono citati esempi di analisi di segnali/immagini (es. rilevazione facciale per dolore) ma senza dettaglio di architetture.
Algoritmo/Modello	Non applicabile; l'articolo è concettuale e non descrive né valida algoritmi specifici.
Compito/Utilizzo	Stratificazione del rischio, previsione in tempo reale di complicanze e outcome (LOS, riammissione), personalizzazione dei piani ERAS (fluidi, analgesia, nutrizione), allocazione efficiente delle risorse e supporto decisionale.
Dataset	Non applicabile; non vengono presentati dataset o analisi originali.
Outcome principali	Non sono riportati outcome quantitativi; il contributo è la mappatura del potenziale ruolo di AI/ML per migliorare aderenza ai protocolli, sicurezza e recupero, oltre alla pianificazione delle risorse.
Metri di valutazione	Non applicabili, in assenza di sperimentazione/validazione empirica.
Risultati chiave	L'IA/ML può abilitare la previsione precoce dei rischi, la personalizzazione degli elementi ERAS e la gestione proattiva delle risorse; gli autori raccomandano un approccio "human-centered" con supervisione clinica, trasparenza e validazioni etiche.
Benefici dichiarati	Potenziale riduzione di complicanze e degenza, miglioramento dell'efficienza e dell'aderenza ai protocolli, percorsi più personalizzati e sicuri per i pazienti.
Limiti dichiarati	Ostacoli a dati condivisibili e di qualità, problemi di privacy e interoperabilità, costi infrastrutturali, gap formativi per i professionisti, rischi di bias e "black box", resistenze organizzative all'adozione.
Riferimento Bibliografico	https://doi.org/10.7759/cureus.56668
Stratificazione per setting	CHIRURGICO
Inclusione studio nella Revisione	SI

ID STUDIO:	ST20
TITOLO:	The Role of Machine Learning in Management of Operating Room: A Systematic Review
Titolo in italiano:	Il ruolo del machine learning nella gestione della sala operatoria: una revisione sistematica.
Autore/i	Alaa Merghani Abdelrazig Merghani; Abdullah Khaled Ahmed Esmail; Ahmed Mohamed Elamin Mubarak Osman; Nihal Ahmed Abdelfrag Mohamed; Safwa Mustafa Mohamed Ali Shentour; Shaima Merghani Abdelrazig Merghani.
Anno	2025
Paese	Sudan, Emirati Arabi Uniti e Arabia Saudita, in base alle affiliazioni riportate.
Tipo di studio	Revisione sistematica condotta secondo le linee guida PRISMA 2020.
Contesto	Chirurgico, con focus sulla gestione organizzativa della sala operatoria (OR) e delle aree perioperatorie.
Specialità/Setting	Gestione trasversale della sala operatoria e PACU in vari ambiti chirurgici (ORL, ortopedia, ginecologia/coloretta, TKA, chirurgia laparoscopica, spinale).

Tecnologia IA	Machine Learning con impiego di metodi di Deep Learning e NLP in alcuni studi (es. Clinical BERT).
Algoritmo/Modello	Panoramica su Random Forest, XGBoost, reti neurali (feed-forward/MLP), SVM, CatBoost, metodi ensemble (bagging/balanced bagging), oltre a modelli NLP (Clinical BERT).
Compito/Utilizzo	Previsione della durata degli interventi, della PACU LOS e delle cancellazioni chirurgiche; ottimizzazione di scheduling, allocazione risorse e flussi perioperatori; supporto decisionale organizzativo.
Dataset	Sintesi di 21 studi inclusi (dal 2019 al 2025) su dati real-world ospedalieri, con campioni da 20 a 124.528 casi; prevalenza di disegni osservazionali retrospettivi e un RCT.
Outcome principali	Dimostrazione che i modelli ML migliorano l'accuratezza di previsione della durata chirurgica e della permanenza in PACU, supportano l'ottimizzazione delle agende/risorse e contribuiscono alla riduzione di ritardi e cancellazioni.
Metri di valutazione	Ampio uso di AUC (circa 0,662–0,82), RMSE, MAE, explained variance e MAPE per stimare performance predittive e scostamenti temporali.
Risultati chiave	Le tecniche ensemble (Random Forest, XGBoost, bagging) sono risultate tra le più performanti; variabili come BMI, sesso, durata pianificata e complessità procedurale emergono come predittori rilevanti; in alcuni studi si osservano riduzioni significative dei tempi di attesa in PACU e miglioramenti dello scheduling.
Benefici dichiarati	Potenziare aumento dell'efficienza operativa, della produttività ospedaliera e della sicurezza, con migliore allocazione delle risorse e coordinamento dei flussi perioperatori.
Limiti dichiarati	Rischio di bias da selezione (inglese-only), eterogeneità metodologica che ostacola meta-analisi, scarsa validazione esterna di molti modelli, e criticità su privacy, accesso ai dati e trasferibilità multi-centro.
Riferimento Bibliografico	https://doi.org/10.7759/cureus.79400
Stratificazione per setting	CHIRURGICO
Inclusione studio nella Revisione	SI

ID STUDIO:	ST21
TITOLO:	Machine Learning for the Prediction of Procedural Case Durations Developed Using a Large Multicenter Database: Algorithm Development and Validation Study
Titolo in italiano:	Machine Learning per la predizione della durata delle procedure chirurgiche sviluppato utilizzando un ampio database multicentrico: studio di sviluppo e validazione dell'algoritmo
Autore/i	Samir Kendale, Andrew Bishara, Michael Burns, Stuart Solomon, Matthew Corriere, Michael Mathis
Anno	2023
Paese	Stati Uniti
Tipo di studio	Studio di sviluppo e validazione algoritmica (osservazionale multicentrico retrospettivo)
Contesto	Chirurgico
Specialità/Setting	Chirurgia in generale, focus sulla predizione della durata delle procedure in sala operatoria

Tecnologia IA	Machine Learning
Algoritmo/Modello	Gradient Boosting, Random Forest, Elastic Net, e modelli ensemble; confrontati con la stima tradizionale dei chirurghi
Compito/Utilizzo	Previsione della durata delle procedure chirurgiche per ottimizzare la gestione delle sale operatorie.
Dataset	Database multicentrico con oltre 120.000 procedure provenienti da più ospedali statunitensi; dati real-world estratti da sistemi informativi ospedalieri.
Outcome principali	Accuratezza predittiva dei modelli ML rispetto alle stime fornite dai chirurghi; confronto tra diversi algoritmi.
Metri di valutazione	MAE (Mean Absolute Error), RMSE (Root Mean Square Error), R ² ; confronto diretto con previsioni umane.
Risultati chiave	I modelli ML hanno sovraperformato le stime dei chirurghi nella previsione della durata delle procedure. Il Gradient Boosting è risultato il migliore, con un errore medio inferiore di diversi minuti rispetto alla stima clinica.
Benefici dichiarati	Miglioramento nella pianificazione delle sale operatorie, riduzione dei ritardi e delle cancellazioni, uso più efficiente delle risorse ospedaliere.
Limiti dichiarati	Mancanza di generalizzabilità a contesti non statunitensi, necessità di validazioni prospettiche, possibile presenza di bias nei dati raccolti retrospettivamente.
Riferimento Bibliografico	https://ai.jmir.org/2023/1/e44909
Stratificazione per setting	CHIRURGICO
Inclusione studio nella Revisione	SI

ID STUDIO:

ST22

TITOLO:

The Application of Artificial Intelligence in Health Care Resource Allocation Before and During the COVID-19 Pandemic: Scoping Review

Titolo in italiano:

L'applicazione dell'intelligenza artificiale nell'allocazione delle risorse sanitarie prima e durante la pandemia di COVID-19: una scoping review

Autore/i

Hao Wu, Xiaoyu Lu, Hanyu Wang

Anno

2023

Paese

Cina (autori affiliati a università cinesi e britanniche)

Tipo di studio

Scoping Review

Contesto

Entrambi; focus trasversale su contesti ospedalieri, di comunità e emergenziali

Specialità/Setting

Gestione delle risorse sanitarie (posti letto, ventilatori, forniture mediche, personale) in setting clinici e di sanità pubblica

Tecnologia IA

Machine Learning, Deep Learning, Reinforcement Learning, NLP

Algoritmo/Modello

Vari modelli identificati negli studi inclusi: Random Forest, Support Vector Machine, Reti Neurali Convoluzionali, modelli di Reinforcement Learning per l'allocazione dinamica, NLP per l'analisi di dati testuali e social media

Compito/Utilizzo

Previsione della domanda di risorse, ottimizzazione dell'allocazione, supporto decisionale per triage e pianificazione sanitaria

Dataset

Dataset eterogenei: cartelle cliniche elettroniche, database amministrativi ospedalieri, dati di monitoraggio COVID-19, social media. Molti real-world, alcuni simulati per scenari di stress delle risorse

Outcome principali

Identificazione delle aree in cui l'IA ha contribuito a migliorare la gestione delle risorse, ridurre sprechi, aumentare l'efficienza e supportare decisioni rapide in emergenza

Metri di valutazione

Varie metriche riportate a seconda degli studi inclusi: AUC >0.85 in modelli

Risultati chiave	<p>predittivi clinici, RMSE ridotti nei modelli di previsione domanda, % di miglioramento nell'allocazione stimata fino al 20–30%</p> <p>L'IA ha mostrato capacità significativa nel predire fabbisogno di risorse, ottimizzare l'uso di posti letto e ventilatori, migliorare la logistica dei dispositivi di protezione individuale e supportare la gestione del personale sanitario. È emersa una forte eterogeneità metodologica tra gli studi</p>
Benefici dichiarati	Migliore resilienza dei sistemi sanitari, riduzione dei tempi di risposta nelle emergenze, maggiore accuratezza nella distribuzione di risorse scarse, supporto all'efficienza ospedaliera
Limiti dichiarati	Mancanza di validazione prospettica, disomogeneità tra studi, problemi di interoperabilità dei dati, rischi di bias algoritmico, mancanza di trasparenza nei modelli adottati
Riferimento Bibliografico	https://doi.org/10.2196/38397
Stratificazione per setting	AMBITO GENERALISTA AI IN SANITA'
Inclusione studio nella Revisione	NO

ID STUDIO: ST23

TITOLO: Artificial intelligence and surgery

Titolo in italiano: Intelligenza artificiale e chirurgia

Autore/i	Masashi Takeuchi, Yuko Kitagawa
Anno	2023
Paese	Giappone (Keio University School of Medicine, Tokyo)
Tipo di studio	Editoriale di revisione/Commentary
Contesto	Chirurgico
Specialità/Setting	Chirurgia gastrointestinale e chirurgia generale, con focus trasversale su pianificazione, esecuzione e follow-up postoperatorio
Tecnologia IA	Intelligenza artificiale con enfasi su Deep Learning, image recognition, analisi video, realtà virtuale/aumentata e sistemi robotici intelligenti
Algoritmo/Modello	Non viene descritto un singolo modello; menzionati approcci di deep learning per riconoscimento anatomico e process recognition in laparoscopia e robotica
Compito/Utilizzo	Supporto decisionale clinico, predizione complicanze, pianificazione preoperatoria, navigazione intraoperatoria, monitoraggio postoperatorio, training chirurgico personalizzato
Dataset	Non applicabile (editoriale narrativo con riferimenti a studi precedenti)
Outcome principali	L'articolo discute i potenziali vantaggi dell'integrazione dell'IA in tutte le fasi della chirurgia: pianificazione (3D imaging, predizione complicanze), esecuzione (robotica e AI navigation), follow-up (monitoraggio complicanze con sensori), formazione (simulazioni AR/VR personalizzate)
Metri di valutazione	Non applicabili (assenza di sperimentazione originale)
Risultati chiave	L'IA è vista come strumento rivoluzionario capace di aumentare la precisione chirurgica, migliorare l'outcome dei pazienti, ridurre complicanze, accelerare il recupero e innovare la formazione dei chirurghi tramite simulazioni immersive
Benefici dichiarati	Maggiore accuratezza chirurgica, riduzione errori, supporto al processo decisionale, percorsi personalizzati, training adattivo per chirurghi in formazione
Limiti dichiarati	Questioni etiche (privacy, sicurezza dati, responsabilità in caso di errori), costi e disuguaglianze di accesso, necessità di formazione specifica per l'uso dei sistemi AI-based

Riferimento Bibliografico	https://doi.org/10.1002/ags3.12766
Stratificazione per setting	CHIRURGICO
Inclusione studio nella Revisione	NO
ID STUDIO:	ST24
TITOLO:	Artificial intelligence-driven surgical innovation: A catalyst for medical equity
Titolo in italiano:	Innovazione chirurgica guidata dall'intelligenza artificiale: un catalizzatore per l'equità sanitaria
Autore/i	Si-Wai Vivian Chiu, Chung-Feng Liu, Kuang-Ming Liao, Chong-Chi Chiu
Anno	2024
Paese	Stati Uniti (Brown University), Taiwan (Chi Mei Medical Center, Min-Hwei Junior College, I-Shou University)
Tipo di studio	Lettera all'editore (commento con revisione narrativa)
Contesto	Chirurgico
Specialità/Setting	Chirurgia generale e chirurgia globale, con focus sui sistemi sanitari dei paesi a basso e medio reddito (LMICs)
Tecnologia IA	Intelligenza artificiale applicata in diversi ambiti: machine learning, computer vision, robotica chirurgica, modelli linguistici di grandi dimensioni (LLM)
Algoritmo/Modello	Non vengono testati modelli specifici; discussione su ML, computer vision e LLM come strumenti chiave
Compito/Utilizzo	Supporto decisionale intraoperatorio, training chirurgico personalizzato, ottimizzazione della gestione delle risorse e riduzione delle disuguaglianze nell'accesso alle cure chirurgiche
Dataset	Non applicabile (commento concettuale senza sperimentazione empirica)
Outcome principali	Discussione del potenziale dell'IA nel migliorare la sicurezza chirurgica, la formazione e soprattutto l'equità nell'accesso alle cure, con focus sulla riduzione delle 18 milioni di morti evitabili annuali nei paesi a basso e medio reddito grazie al rafforzamento della chirurgia globale
Metri di valutazione	Non applicabili (assenza di studio sperimentale)
Risultati chiave	L'IA può rafforzare la valutazione del rischio, migliorare il monitoraggio postoperatorio, ottimizzare la logistica e fornire supporto intraoperatorio basato su evidenze. L'integrazione di LLM e computer vision offre nuove opportunità per formazione, simulazione e accesso a linee guida cliniche
Benefici dichiarati	Potenziale riduzione delle disparità globali nell'accesso alla chirurgia, miglioramento degli outcome, formazione chirurgica più accessibile e personalizzata, sistemi sanitari più resilienti
Limiti dichiarati	Sfide legate alla governance dei dati, alla sicurezza, alla rappresentatività dei dataset, alle barriere etiche e culturali, e alla resistenza da parte degli operatori sanitari
Riferimento Bibliografico	https://doi.org/10.1002/ags3.12827
Stratificazione per setting	CHIRURGICO
Inclusione studio nella Revisione	NO

ID STUDIO: ST25
TITOLO: A Review on the Current Applications of Artificial Intelligence in the Operating Room
Titolo in italiano: Una revisione sulle applicazioni attuali dell'intelligenza artificiale in sala operatoria

Autore/i	David C. Birkhoff, Anne Sophie H.M. van Dalen, Marlies P. Schijven
Anno	2021
Paese	Paesi Bassi (Università di Amsterdam)
Tipo di studio	Revisione sistematica
Contesto	Chirurgico
Specialità/Setting	Sala operatoria (chirurgia generale e specialità varie)
Tecnologia IA	Machine Learning, Deep Learning, Computer Vision
Algoritmo/Modello	Artificial Neural Networks (ANN), Convolutional Neural Networks (CNN), Support Vector Machine (SVM), Naïve Bayes, Recurrent Neural Networks (RNN), Adaptive Neuro-Fuzzy Inference System (ANFIS)
Compito/Utilizzo	Previsione della durata degli interventi, riconoscimento gestuale e delle fasi chirurgiche, rilevamento strumenti, guida autonoma di endoscopi e robot chirurgici, supporto al training
Dataset	9 studi inclusi, con dataset eterogenei (registrazioni video real-world da procedure chirurgiche, simulatori, database di movimenti e strumenti)
Outcome principali	Le applicazioni AI in sala operatoria sono già realtà in diversi ambiti: predizione di durata, riconoscimento di gesti e fasi chirurgiche, automazione di strumenti, monitoraggio operatori e supporto didattico
Metri di valutazione	Variano a seconda degli studi: accuratezza del riconoscimento (fino al 96%), AUC, errori medi assoluti ridotti per la durata chirurgica, precision e recall nei compiti di computer vision
Risultati chiave	L'IA ha dimostrato ottime prestazioni in compiti di riconoscimento visivo e predizione; le CNN e le RNN hanno raggiunto elevata accuratezza per classificazione delle fasi chirurgiche, mentre modelli ensemble hanno migliorato la predizione dei tempi operatori. Le applicazioni sono promettenti anche per la formazione chirurgica
Benefici dichiarati	Maggiore efficienza in sala operatoria, ottimizzazione delle risorse, supporto decisionale, possibilità di training automatizzato per chirurghi in formazione
Limiti dichiarati	Numero ridotto di studi disponibili (solo 9), eterogeneità di algoritmi e metriche, mancanza di validazioni prospettiche multicentriche, rischio di scarsa trasferibilità dei modelli alla pratica clinica
Riferimento Bibliografico	https://doi.org/10.1177/1553350621996961
Stratificazione per setting	CHIRURGICO
Inclusione studio nella Revisione	SI

ID STUDIO:	ST26
TITOLO:	The Legal Governance of AI Around the World
Titolo in italiano:	La governance legale dell'intelligenza artificiale nel mondo
Autore/i	Effy Vayena, Alessandro Blasimme, I. Glenn Cohen
Anno	2023
Paese	Internazionale (analisi comparata di diversi ordinamenti: USA, UE, Cina, Regno Unito, Canada, Singapore, Giappone, Corea del Sud)
Tipo di studio	Revisione narrativa / Analisi comparativa di tipo giuridico
Contesto	Non applicabile (focus normativo e regolatorio)
Specialità/Setting	Regolamentazione dell'intelligenza artificiale in sanità e in altri ambiti ad alto impatto (sanità, sicurezza, giustizia, infrastrutture critiche)
Tecnologia IA	Tutte le principali tecnologie IA considerate in ottica regolatoria (ML, DL, NLP, sistemi esperti, AI generativa)
Algoritmo/Modello	Non applicabile; focus su framework legali, non su validazione di modelli
Compito/Utilizzo	Non applicabile in senso tecnico; l'IA è analizzata come oggetto di governance legale per mitigare rischi e garantire uso etico e sicuro
Dataset	Non applicabile (non è uno studio empirico ma un'analisi normativa)
Outcome principali	Identificazione delle differenze e similitudini tra i quadri normativi globali; mappatura dei modelli di governance emergenti (basati su risk-based approach, responsabilità legale, etica e trasparenza)
Metri di valutazione	Non applicabili Risultati chiave: La regolamentazione dell'IA sta convergendo su approcci risk-based (UE, Canada, Singapore) e su normative settoriali (USA, UK, Giappone). Cina e Corea del Sud hanno introdotto quadri più stringenti sul controllo statale e sulla supervisione etica. Permangono differenze sostanziali su responsabilità civile, trasparenza algoritmica, certificazioni e diritti degli utenti
Risultati chiave	Non applicabili
Benefici dichiarati	Maggiore chiarezza regolatoria, tutela dei cittadini, promozione di un'IA etica e affidabile, creazione di condizioni per un'innovazione responsabile
Limiti dichiarati	Frammentazione normativa, mancanza di standard internazionali condivisi, rischi di conflitti normativi transfrontalieri, difficoltà di adattamento della regolazione a un'innovazione molto rapida
Riferimento Bibliografico	https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4349661
Stratificazione per setting	AMBITO GENERALISTA AI IN SANITA'
Inclusione studio nella Revisione	NO

ID STUDIO:	ST27
TITOLO:	Natural Language Processing (NLP)- and Machine Learning (ML)-Enabled Operating Room Optimization: A Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) Systematic Review Anchored in Project Planning Theory
Titolo in italiano:	Ottimizzazione della sala operatoria abilitata da NLP (Natural Language Processing) e ML (Machine Learning): revisione sistematica secondo PRISMA basata sulla teoria della pianificazione dei progetti

Autore/i	Balaiah Chamarthi, Omkar Reddy Polu, Sathish Krishna Anumula, Azhar Ushmani, Pratik Kasralikar, Abdul Aleem Syed
Anno	2023
Paese	India e USA (in base alle affiliazioni riportate)
Tipo di studio	Revisione sistematica (PRISMA)
Contesto	Chirurgico
Specialità/Setting	Gestione e organizzazione della sala operatoria (diverse specialità chirurgiche)
Tecnologia IA	Machine Learning e Natural Language Processing
Algoritmo/Modello	Random Forest, Support Vector Machines (SVM), XGBoost, reti neurali profonde, modelli NLP come BERT; usati nei vari studi inclusi per analizzare dati clinici e note testuali
Compito/Utilizzo	Previsione durata degli interventi, ottimizzazione scheduling chirurgico, analisi flussi operatori, supporto decisionale per gestione risorse
Dataset	Studi inclusi con dataset eterogenei (EHR real-world, note operatorie testuali, dataset amministrativi ospedalieri); dimensioni variabili da centinaia a oltre 100.000 procedure
Outcome principali	L'IA, in particolare ML e NLP, migliora la capacità predittiva e ottimizza la gestione delle risorse operatorie; identificati benefici concreti su scheduling, riduzione dei ritardi, analisi automatizzata dei report testuali
Metri di valutazione	MAE, RMSE, AUC, precision, recall, explained variance, MAPE; risultati variabili ma in genere superiori ai metodi tradizionali
Risultati chiave	I modelli ML hanno mostrato performance elevate rispetto alle stime manuali; l'integrazione NLP consente di estrarre informazioni da note cliniche non strutturate, arricchendo la predittività; il framework PRISMA ha permesso di sintetizzare evidenze trasversali
Benefici dichiarati	Migliore efficienza della sala operatoria, riduzione tempi morti e cancellazioni, ottimizzazione delle risorse, maggiore accuratezza predittiva e analisi automatizzata della documentazione clinica
Limiti dichiarati	Eterogeneità degli studi inclusi, mancanza di validazioni prospettiche su larga scala, potenziali bias nei dataset, necessità di standardizzazione degli approcci NLP/ML in ambito chirurgico
Riferimento Bibliografico	https://doi.org/10.7759/cureus.82796
Stratificazione per setting	CHIRURGICO
Inclusione studio nella Revisione	SI

ID STUDIO:	ST28
TITOLO:	Inteligencia Artificial en la gestión de los servicios de salud: Estado actual y perspectivas futuras
Titolo in italiano:	Intelligenza artificiale nella gestione dei servizi sanitari: stato attuale e prospettive future

Autore/i	Dennys Raquel Ortiz Luzuriaga, Víctor E. Briones Morales
Anno	2025
Paese	Ecuador
Tipo di studio	Revisione narrativa / studio di prospettiva
Contesto	Entrambi, con focus sulla gestione dei servizi sanitari in senso ampio
Specialità/Setting	Gestione dei servizi sanitari, amministrazione ospedaliera, processi clinico-

Tecnologia IA	assistenziali e sanitari pubblici Machine Learning, Deep Learning, Natural Language Processing, Computer Vision, Robotics, sistemi esperti
Algoritmo/Modello	Non sono descritti algoritmi specifici, ma vengono discussi i principali approcci ML/DL applicabili ai servizi sanitari
Compito/Utilizzo	Previsione di domanda di servizi, ottimizzazione della gestione delle risorse, supporto decisionale clinico e amministrativo, automazione di processi assistenziali e gestionali
Dataset	Non vengono analizzati dataset specifici; discussione generale con esempi da letteratura e casi applicativi real-world riportati in studi precedenti
Outcome principali	L'IA ha già un impatto crescente sulla gestione dei servizi sanitari migliorando l'efficienza organizzativa, la precisione diagnostica e la qualità assistenziale; le prospettive future vedono un'integrazione ancora più estesa nei processi decisionali e gestionali
Metri di valutazione	Non applicabili (assenza di dati quantitativi originali)
Risultati chiave	L'IA è considerata uno strumento strategico per la trasformazione digitale della sanità, con potenziale per ottimizzare risorse, migliorare outcome clinici e ridurre costi; si sottolinea l'importanza di un quadro normativo chiaro e di un approccio etico nell'adozione
Benefici dichiarati	Efficienza, accuratezza, riduzione errori umani, supporto alla gestione integrata e maggiore accessibilità ai servizi sanitari
Limiti dichiarati	Mancanza di standardizzazione, rischi etici e di privacy, dipendenza da infrastrutture tecnologiche, necessità di formazione degli operatori
Riferimento Bibliografico	https://doi.org/10.70577/ASCE/483.502/2025
Stratificazione per setting	AMBITO GENERALISTA AI IN SANITA'
Inclusione studio nella Revisione	SI

ID STUDIO:	ST29
TITOLO:	Predicting Robotic Hysterectomy Incision Time: Optimizing Surgical Scheduling with Machine Learning
Titolo in italiano:	Predire i tempi di incisione nell'isterectomia robotica: ottimizzazione della programmazione chirurgica con il machine learning
Autore/i	Vaishali Shah, Halley C. Yung, Jie Yang, Justin Zaslavsky, Gabriela N. Algarroba, Alyssa Pullano, Hannah C. Karpel, Nicole Munoz, Yindalon Aphinyanaphongs, Mark Saraceni, Paresh Shah, Simon Jones, Kathy Huang
Anno	2024
Paese	Stati Uniti
Tipo di studio	Studio retrospettivo osservazionale di sviluppo e validazione algoritmica
Contesto	Chirurgico
Specialità/Setting	Ginecologia robotica (isterectomia)
Tecnologia IA	Machine Learning
Algoritmo/Modello	Linear Ridge Regression, Random Forest, XGBoost, CatBoost, Explainable Boosting Machine (EBM)
Compito/Utilizzo	Previsione della durata dei tempi di incisione chirurgica per ottimizzare la programmazione delle sale operatorie
Dataset	3.058 casi di isterectomia robotica identificati dal database EMR (Electronic Medical Record), ridotti a 2.702 dopo criteri di selezione (chirurghi con >50 casi;

Outcome principali	incisioni <3 ore). Fonte: dati real-world multicentrici
Metri di valutazione	accuratezza predittiva dei modelli ML rispetto alla baseline (stime manuali basate sulla media dei tempi precedenti del chirurgo)
Risultati chiave	Accuratezza entro $\pm 15\%$ del tempo reale, errori medi assoluti, wait-time blocks, excess-time blocks
Benefici dichiarati	L'Explainable Boosting Machine (EBM) ha ottenuto i risultati migliori, con una riduzione del 21% degli excess-time blocks (da 1.113 a 905), pari a circa 52 ore di tempo operativo recuperato. Tutti i modelli ML hanno superato le stime tradizionali dei chirurghi
Limiti dichiarati	Migliore efficienza della sala operatoria, riduzione dei tempi morti, maggiore prevedibilità, incremento del throughput chirurgico e dei ricavi ospedalieri
Riferimento Bibliografico	Mancata generalizzabilità ad altri contesti, necessità di revisione manuale di alcune variabili (es. dimensioni uterine), esclusione dei casi >3 ore
Stratificazione per setting	https://doi.org/10.4293/JSLS.2024.00040
Inclusione studio nella Revisione	CHIRURGICO SI

ID STUDIO: ST30

TITOLO: Artificial Intelligence for Patient Flow

Titolo in italiano: Intelligenza artificiale per il flusso dei pazienti

Autore/i	CADTH (Canadian Agency for Drugs and Technologies in Health)
Anno	2024
Paese	Canada
Tipo di studio	Horizon Scan (revisione narrativa di evidenze emergenti)
Contesto	Entrambi, con prevalenza ospedaliera (PS, reparti medici, oncologia, sale operatorie)
Specialità/Setting	Gestione del patient flow ospedaliero e territoriale
Tecnologia IA	Machine Learning, modelli predittivi, NLP in alcuni progetti
Algoritmo/Modello	Non vengono indicati modelli unici; menzionati strumenti predittivi per ammissioni, dimissioni, riammissioni e scheduling (es. Signal 1 in Canada)
Compito/Utilizzo	Non applicabile in senso tecnico; l'IA è analizzata come oggetto di governance legale per mitigare rischi e garantire uso etico e sicuro
Dataset	Non applicabile (non è uno studio empirico ma un'analisi normativa)
Outcome principali	Identificazione delle differenze e similitudini tra i quadri normativi globali; mappatura dei modelli di governance emergenti (basati su risk-based approach, responsabilità legale, etica e trasparenza)
Metri di valutazione	Non applicabili Risultati chiave: La regolamentazione dell'IA sta convergendo su approcci risk-based (UE, Canada, Singapore) e su normative settoriali (USA, UK, Giappone). Cina e Corea del Sud hanno introdotto quadri più stringenti sul controllo statale e sulla supervisione etica. Permangono differenze sostanziali su responsabilità civile, trasparenza algoritmica, certificazioni e diritti degli utenti
Risultati chiave	Non applicabili
Benefici dichiarati	Maggiore chiarezza regolatoria, tutela dei cittadini, promozione di un'IA etica e affidabile, creazione di condizioni per un'innovazione responsabile
Limiti dichiarati	Frammentazione normativa, mancanza di standard internazionali condivisi, rischi di conflitti normativi transfrontalieri, difficoltà di adattamento della regolazione a

**Riferimento
Bibliografico
Stratificazione per
setting
Inclusione studio
nella Revisione**

un'innovazione molto rapida

<https://www.ncbi.nlm.nih.gov/books/n/op728369720125/pdf/>

AMBITO GENERALISTA AI IN SANITA'

NO

ID STUDIO:

ST31

TITOLO:

Enhancing Operating Room Surgical Efficiency through Artificial Intelligence: A Comprehensive Review

Titolo in italiano:

Migliorare l'efficienza chirurgica in sala operatoria attraverso l'intelligenza artificiale: una revisione completa

Autore/i

Alejandro Espallat

Anno

2024

Paese

Stati Uniti

Tipo di studio

Revisione narrativa

Contesto

Chirurgico

Specialità/Setting

Sala operatoria, chirurgia multidisciplinare

Tecnologia IA

Machine Learning, Deep Learning, Computer Vision, NLP, sistemi ibridi, Augmented Reality (AR)

Algoritmo/Modello

Convolutional Neural Networks (CNN), Random Forest, XGBoost, Artificial Neural Networks (ANN), Explainable AI, strumenti di Surgical Data Science (es. OR-Tools di Google, script Python di scheduling)

Compito/Utilizzo

Previsione della durata operatoria, ottimizzazione dello scheduling, supporto decisionale intraoperatorio, automazione robotica, valutazione delle abilità chirurgiche

Dataset

Diversi dataset provenienti da registri ospedalieri real-world, Electronic Health Records (EHR), database chirurgici (es. NSQIP), simulazioni AR e dati intraoperatori (centinaia a decine di migliaia di casi a seconda dello studio incluso)

Outcome principali

Riduzione di ritardi e straordinari, miglior utilizzo della sala operatoria, predizione accurata dei tempi chirurgici, identificazione precoce di complicanze, miglioramento della formazione chirurgica

Metri di valutazione

AUC fino a 0,87 per predizione riammissioni, miglioramento dell'accuratezza predittiva fino al 16,8%, riduzione del 21% degli straordinari infermieristici, risparmi stimati di circa 469.000 \$ in 3 anni

Risultati chiave

L'IA ha mostrato forte impatto nella pianificazione chirurgica e nell'efficienza intraoperatoria, superando i metodi tradizionali. Applicazioni di AR e robotica aumentano precisione e training. L'uso di ML ha permesso di ottimizzare scheduling e predizione complicanze

Benefici dichiarati

Efficienza, riduzione costi, maggiore sicurezza, supporto decisionale in tempo reale, miglioramento della qualità assistenziale, training chirurgico avanzato

Limiti dichiarati

Eterogeneità dei dati, necessità di dataset ampi e annotati, scarsa validazione multicentrica, difficoltà di integrazione clinica, questioni etiche e regolatorie, rischio di bias algoritmico

Riferimento

<https://doi.org/10.33425/2689-1093.1073>

Bibliografico

Stratificazione per

setting

CHIRURGICO

Inclusione studio
nella Revisione

SI

ID STUDIO:

ST32

TITOLO:

Real-Time Analytics and AI for Managing No-Show Appointments in Primary Health Care in the United Arab Emirates: Before-and-After Study

Titolo in italiano:

Analitiche in tempo reale e intelligenza artificiale per la gestione delle mancate presentazioni agli appuntamenti nelle cure primarie negli Emirati Arabi Uniti: studio prima-dopo

Autore/i

Yousif Mohamed ALSerkal, Naseem Mohamed Ibrahim, Aisha Suhail Alsereidi, Mubarak Ibrahim, Sudheer Kurakula, Sadaf Ahsan Naqvi, Yasir Khan, Neema Preman Oottumadathil

Anno

2025

Paese

Emirati Arabi Uniti

Tipo di studio

Studio osservazionale before-and-after

Contesto

Ambulatoriale (Primary Health Care)

Specialità/Setting

Cure primarie – gestione appuntamenti e pazienti

Tecnologia IA

Machine Learning

Algoritmo/Modello

Random Forest Classifier (accuratezza 86%)

Compito/Utilizzo

Previsione rischio no-show e ottimizzazione della gestione delle agende tramite dashboard in tempo reale

Dataset

135.393 appuntamenti (67.429 pre-implementazione, 67.964 post-implementazione) da dati real-world dell'Emirates Health Services

Outcome principali

Riduzione del tasso di no-show e dei tempi medi di attesa dei pazienti

Metri di valutazione

Accuracy 86%; Odds ratio 0,43; riduzione no-show del 50,7%; riduzione media dei tempi di attesa di 5,7 minuti

Risultati chiave

I no-show sono passati dal 20,82% al 10,25%; i tempi di attesa medi si sono ridotti da 54,2 a 48,5 minuti, con miglioramenti fino al 50% in alcuni centri; oltre 6.400 ore di tempo clinico risparmiate in 3 mesi

Benefici dichiarati

Migliore efficienza organizzativa, riduzione costi, maggiore soddisfazione del paziente, ottimizzazione uso risorse e pianificazione più accurata

Limiti dichiarati

Possibile variabilità nei centri, necessità di validazioni prospettiche su larga scala, dipendenza dalla qualità dei dati raccolti nell'EHR

Riferimento

<https://doi.org/10.2196/64936>

Bibliografico

Stratificazione per
setting

AMBULATORIALE

Inclusione studio
nella Revisione

SI

ID STUDIO:

ST33

TITOLO:

First deployment of artificial intelligence recommendations in orthopedic surgery

Titolo in italiano:

Prima implementazione di raccomandazioni basate su intelligenza

artificiale in chirurgia ortopedica

Autore/i	Farid Al Zoubi, Koorosh Kashanian, Paul Beaulé, Pascal Fallavollita
Anno	2024
Paese	Canada
Tipo di studio	Community Case Study (studio osservazionale applicativo)
Contesto	Chirurgico
Specialità/Setting	Chirurgia ortopedica (artroplastica anca e ginocchio)
Tecnologia IA	Machine Learning – prescriptive analytics
Algoritmo/Modello	Decision tree e modelli interpretabili per benchmark predittivi
Compito/Utilizzo	Supporto decisionale e ottimizzazione dell'efficienza operatoria
Dataset	228 interventi di artroplastica (anca/ginocchio) eseguiti al The Ottawa Hospital, 2023; real-world data; coinvolti 7 chirurghi, 43 infermieri, 13 anestesisti
Outcome principali	Surgical Success Rate (SSR), definito come completamento di 4 interventi in 8 ore
Metri di valutazione	Surgical Success Rate (SSR) in %; confronto pre e post implementazione PAS
Risultati chiave	SSR migliorato dal 39% al 93%; in oltre il 50% dei casi è stato possibile aggiungere un 5° intervento nella stessa giornata; riduzione significativa dei tempi morti
Benefici dichiarati	Aumento produttività, riduzione costi da straordinari (~570.000 \$/anno), ottimizzazione flussi operatori, maggiore capacità trattare pazienti nello stesso arco temporale
Limiti dichiarati	Applicazione limitata ad artroplastica ad alto volume; implementazione manuale con benchmark fissi; possibile bias del giorno (sabato); differenze di genere nei tempi operatori; assenza di implementazione real-time
Riferimento Bibliografico	https://doi.org/10.3389/frai.2024.1342234
Stratificazione per setting	CHIRURGICO
Inclusione studio nella Revisione	SI

ID STUDIO:	ST34
TITOLO:	A Machine Learning-Based Approach for Predicting Patient Punctuality in Ambulatory Care Centers
Titolo in italiano:	Un approccio basato sul machine learning per predire la puntualità dei pazienti nei centri di cure ambulatoriali

Autore/i	Sharan Srinivas
Anno	2020
Paese	Stati Uniti
Tipo di studio	Studio metodologico osservazionale con applicazione di ML su dati real-world
Contesto	Ambulatoriale
Specialità/Setting	Centri di cure ambulatoriali (multi-specialità, scheduling e gestione flussi)
Tecnologia IA	Machine Learning
Algoritmo/Modello	Decision Tree, Random Forest, Gradient Boosting, Support Vector Machines (SVM), k-Nearest Neighbors (k-NN), Logistic Regression, Naïve Bayes
Compito/Utilizzo	Previsione della puntualità dei pazienti agli appuntamenti ambulatoriali
Dataset	57.000 appuntamenti raccolti in 3 anni da un centro ambulatoriale negli Stati Uniti (dataset real-world)
Outcome principali	Accuratezza predittiva dei modelli ML per classificare i pazienti come puntuali o in ritardo (>15 minuti)

Metri di valutazione	Accuracy, Precision, Recall, F1-score, AUC. Random Forest ha ottenuto AUC = 0,84, F1-score = 0,79
Risultati chiave	Random Forest e Gradient Boosting sono risultati i migliori algoritmi. Fattori più predittivi: orario dell'appuntamento, giorno della settimana, storico di puntualità del paziente, distanza geografica. I modelli ML superano nettamente le previsioni tradizionali (basate su medie storiche semplici)
Benefici dichiarati	Possibilità di ridurre tempi morti, ottimizzare agende e risorse, migliorare la gestione dei flussi ambulatoriali, aumentare soddisfazione pazienti e staff
Limiti dichiarati	Studio condotto in un solo centro ambulatoriale → limitata generalizzabilità; mancano fattori socio-demografici e psicologici; necessità di validazioni esterne in contesti diversi
Riferimento Bibliografico	https://doi.org/10.3390/ijerph17103703
Stratificazione per setting	AMBULATORIALE
Inclusione studio nella Revisione	SI

ID STUDIO: ST35

TITOLO: Artificial Intelligence and Surgical Decision-making

Titolo in italiano: Intelligenza artificiale e processo decisionale chirurgico

Autore/i	Tyler J. Loftus, Patrick J. Tighe, Amanda C. Filiberto, Philip A. Efron, Scott C. Brakenridge, Alicia M. Mohr, Parisa Rashidi, Gilbert R. Upchurch Jr, Azra Bihorac
Anno	2020
Paese	Stati Uniti
Tipo di studio	Revisione narrativa (commentary/review)
Contesto	Chirurgico
Specialità/Setting	Chirurgia generale e sottospecialità (decision-making perioperatorio)
Tecnologia IA	Machine Learning, Deep Learning, Natural Language Processing, Reinforcement Learning
Algoritmo/Modello	Reti neurali profonde, modelli NLP, sistemi di apprendimento per rinforzo per decision-making dinamico (es. gestione fluidi, analgesia, terapia intensiva chirurgica)
Compito/Utilizzo	Supporto decisionale clinico e chirurgico, previsione di complicanze, ottimizzazione trattamenti personalizzati, triage e gestione perioperatoria
Dataset	Vari studi citati, basati su grandi database chirurgici e dati EHR real-world; non un dataset unico, ma revisione di più applicazioni (es. predizione complicanze da NSQIP, analisi immagini radiologiche, note cliniche testuali)
Outcome principali	L'IA ha dimostrato potenziale nel migliorare la precisione predittiva di complicanze, mortalità e outcomes chirurgici rispetto ai sistemi di scoring clinico tradizionali
Metri di valutazione	Reportati AUC superiori rispetto agli score tradizionali (es. AUC >0,90 per predizione mortalità vs score clinici standard); metriche variabili a seconda degli studi riportati
Risultati chiave	L'IA può superare i sistemi tradizionali nella predizione degli esiti post-operatori e fornire supporto decisionale tempestivo e personalizzato; esempi pratici includono predizione sepsi, ARDS, necessità trasfusioni, e gestione fluidi intraoperatoria con RL
Benefici dichiarati	Supporto clinico data-driven, maggiore accuratezza predittiva, personalizzazione delle cure, potenziale riduzione di errori clinici e miglioramento outcomes

Limiti dichiarati	Necessità di grandi dataset annotati, rischio di bias e opacità algoritmica (black-box), mancanza di validazioni multicentriche, scarsa integrazione nei flussi clinici reali, problemi etici e legali sulla responsabilità
Riferimento Bibliografico	https://doi.org/10.1001/jamasurg.2019.4917
Stratificazione per setting	CHIRURGICO
Inclusione studio nella Revisione	SI

ID STUDIO:	ST36
TITOLO:	The Transformative Role of Artificial Intelligence in Plastic and Reconstructive Surgery: Challenges and Opportunities
Titolo in italiano:	Il ruolo trasformativo dell'intelligenza artificiale nella chirurgia plastica e ricostruttiva: sfide e opportunità

Autore/i	Masab Mansoor, Andrew F. Ibrahim
Anno	2025
Paese	Stati Uniti
Tipo di studio	Revisione narrativa / commentary
Contesto	Chirurgico
Specialità/Setting	Chirurgia plastica e ricostruttiva
Tecnologia IA	Machine Learning, Deep Learning, Computer Vision, Natural Language Processing
Algoritmo/Modello	Convolutional Neural Networks (CNN) per analisi di immagini, modelli NLP per analisi cartelle cliniche e note operative; non riportato un unico algoritmo predominante
Compito/Utilizzo	Supporto decisionale, previsione esiti chirurgici, analisi di immagini pre e post-operatorie, personalizzazione piani terapeutici, miglioramento training chirurgico
Dataset	Dati clinici e imaging real-world da cartelle elettroniche e database fotografici di chirurgia plastica; non riportata una dimensione univoca poiché trattasi di revisione
Outcome principali	Potenziabile miglioramento nell'accuratezza diagnostica, previsione delle complicanze e valutazione degli outcome estetici e funzionali; aumento efficienza nel flusso chirurgico e nella pianificazione
Metri di valutazione	Nei lavori citati sono riportati AUC e accuratèzze superiori a quelle dei metodi tradizionali, ma senza valori numerici consolidati all'interno di questa revisione
Risultati chiave	L'IA ha il potenziale di rivoluzionare la chirurgia plastica, supportando la selezione dei candidati, la pianificazione degli interventi e il follow-up; consente di ridurre la variabilità soggettiva nella valutazione clinica e migliorare l'educazione chirurgica
Benefici dichiarati	Migliore precisione clinica, supporto alle decisioni chirurgiche, riduzione complicanze, possibilità di outcome più predittivi e standardizzati, personalizzazione del percorso chirurgico
Limiti dichiarati	Mancanza di dataset ampi e annotati, rischio bias algoritmico, difficoltà di integrazione nei flussi clinici, barriere etico-legali e necessità di validazioni multicentriche
Riferimento Bibliografico	https://doi.org/10.3390/jcm14082698
Stratificazione per setting	CHIRURGICO

Inclusione studio
nella Revisione

SI

ID STUDIO:

ST37

TITOLO:

Applied Artificial Intelligence in Healthcare: A Review of Computer Vision Technology Application in Hospital Settings

Titolo in italiano:

Intelligenza artificiale applicata in sanità: revisione dell'applicazione della Computer Vision in ambito ospedaliero

Autore/i

Heidi Lindroth; Keivan Nalaie; Roshini Raghu; Ivan N. Ayala; Charles Busch; Anirban Bhattacharyya; Pablo Moreno Franco; Daniel A. Diedrich; Brian W. Pickering; Vitaly Herasevich

Anno

2024

Paese

USA (affiliazioni principali: Mayo Clinic; Indiana University; Univ. of Wisconsin)

Tipo di studio

Revisione narrativa/rapida (Review)

Contesto

Ospedaliero, con applicazioni in terapia intensiva, sala operatoria e reparti, ed estensioni in ambito ambulatoriale e comunitario

Specialità/Setting

Terapia intensiva, chirurgia, radiologia, istopatologia, pediatria e neonatologia, riabilitazione e prevenzione delle infezioni

Tecnologia IA

Computer vision basata su machine learning e deep learning, inclusi modelli transformer

Algoritmo/Modello

Convolutional neural networks e varianti come Mask-RCNN, YOLO e U-Net, vision transformers, support vector machine, random forest, k-nearest neighbors, reti ricorrenti LSTM e GRU, regressione logistica e metodi tradizionali di classificazione

Compito/Utilizzo

Rilevazione e monitoraggio di dolore, sedazione, delirium, deterioramento clinico, cadute e igiene delle mani, classificazione e segmentazione di immagini radiologiche e istologiche, supporto decisionale intraoperatorio, quantificazione del carico di lavoro e autodocumentazione clinica

Dataset

Utilizzo di dataset generici come Pascal VOC e COCO e di dataset specifici sanitari quali MCIndoor20000, MYNursingHome e HIOD, oltre a raccolte oncologiche; evidenziata la scarsità di dataset pubblici real-world in ospedale e l'importanza di transfer learning, federated learning e data augmentation

Outcome principali

Applicazioni di computer vision in ospedale con dimostrazione di efficacia in diversi domini, tra cui rilevazione dell'igiene delle mani con sensibilità e specificità superiori al 90%, monitoraggio della mobilità in terapia intensiva con AUROC pari a 0,938 e sensibilità e specificità oltre l'87%, valutazione del dolore neonatale con correlazioni maggiori di 0,8 con le scale cliniche, identificazione precoce del deterioramento clinico con AUC di 0,81 e monitoraggio di cadute e workload con accuratezza e correlazioni elevate

Metri di valutazione

AUC, AUROC, sensibilità, specificità, accuratezza, coefficienti di correlazione, indici di concordanza ICC e kappa ponderato

Risultati chiave

Le applicazioni ospedaliere sono prevalentemente in fase di prototipo o proof-of-concept, mentre risultano più mature le implementazioni in radiologia e istopatologia; principali barriere all'adozione sono privacy, bias dei dati, limitata disponibilità di dataset annotati e difficoltà di integrazione nei flussi clinici complessi

Benefici dichiarati

Monitoraggio continuo e oggettivo, rilevazione precoce del deterioramento e del dolore, miglioramento della sicurezza del paziente tramite prevenzione di cadute e controllo dell'igiene delle mani, riduzione del carico documentale, supporto all'efficienza organizzativa in terapia intensiva e in sala operatoria

Limiti dichiarati

Problemi di privacy e consenso, gestione e archiviazione sicura dei dati video,

**Riferimento
Bibliografico
Stratificazione per
setting
Inclusione studio
nella Revisione**

scarsa generalizzabilità dovuta a dataset poco rappresentativi, rischio di bias, mancanza di validazioni prospettiche e difficoltà di implementazione pratica, elevati requisiti computazionali

<https://doi.org/10.3390/jimaging10040081>

AMBITO GENERALISTA AI IN SANITA'

NO

ID STUDIO:

ST38

TITOLO:

Role of Artificial Intelligence Applications in Real-Life Clinical Practice: Systematic Review

Titolo in italiano:

Ruolo delle applicazioni di intelligenza artificiale nella pratica clinica reale: revisione sistematica

Autore/i

Jiamin Yin, Kee Yuan Ngiam, Hock Hai Teo

Anno

2021

Paese

Singapore

Tipo di studio

Revisione sistematica

Contesto

Prevalentemente ospedaliero con applicazioni anche in ambito ambulatoriale, considerando diversi setting clinici reali

Specialità/Setting

Radiologia, cardiologia, oncologia, dermatologia, oftalmologia, medicina d'urgenza e altri ambiti multidisciplinari

Tecnologia IA

Machine learning e deep learning, con applicazioni legate a modelli predittivi e sistemi di supporto decisionale

Algoritmo/Modello

Convolutional neural networks, support vector machine, random forest, gradient boosting, modelli di deep learning per analisi di immagini e segnali clinici

Compito/Utilizzo

Previsione di esiti clinici, ottimizzazione dei percorsi di diagnosi e trattamento, supporto decisionale per la diagnosi e la gestione terapeutica

Dataset

Utilizzo di dataset clinici real-world provenienti da ospedali e istituzioni sanitarie, con dimensioni variabili da piccoli campioni locali a grandi database multicentrici; prevalenza di dati retrospettivi con limitata disponibilità di trial prospettici

Outcome principali

Identificazione delle applicazioni di IA che hanno raggiunto la fase di implementazione clinica, analisi della loro efficacia e delle aree di utilizzo; maggiore applicazione in radiologia, cardiologia e oncologia; risultati promettenti ma con limitazioni metodologiche e di generalizzabilità

Metri di valutazione

Accuratezza diagnostica, sensibilità, specificità, AUC, tempo di risposta e confronto con performance cliniche standard

Risultati chiave

Le applicazioni di IA mostrano performance comparabili o superiori agli specialisti in determinati ambiti, in particolare nell'analisi delle immagini diagnostiche e nella predizione di esiti clinici. Tuttavia, la maggior parte degli studi è ancora limitata da contesti controllati, con pochi esempi di adozione su larga scala nella pratica clinica quotidiana.

Benefici dichiarati

Maggior accuratezza diagnostica, rapidità nell'interpretazione dei dati, supporto decisionale utile per migliorare la gestione clinica e potenziale riduzione dei tempi e dei costi sanitari

Limiti dichiarati

Mancanza di validazioni prospettiche robuste, limitata generalizzabilità dei modelli sviluppati su singoli dataset, problematiche etiche e di privacy, difficoltà di integrazione con i flussi di lavoro clinici esistenti

Riferimento

<https://doi.org/10.2196/25759>

Bibliografico
Stratificazione per setting
Inclusione studio nella Revisione

AMBITO GENERALISTA AI IN SANITA'

NO

ID STUDIO:

ST39

TITOLO:

Effects of an Artificial Intelligence Platform for Behavioral Interventions on Depression and Anxiety Symptoms: Randomized Clinical Trial

Titolo in italiano:

Effetti di una piattaforma di intelligenza artificiale per interventi comportamentali sui sintomi di depressione e ansia: trial clinico randomizzato

Autore/i

Shiri Sadeh-Sharvit, T Del Camp, Sarah E Horton, Jacob D Hefner, Jennifer M Berry, Eyal Grossman, Steven D Hollon

Anno

2023

Paese

Stati Uniti

Tipo di studio

Randomized Clinical Trial (RCT)

Contesto

Ambulatoriale (terapia cognitivo-comportamentale in clinica comunitaria)

Specialità/Setting

Salute mentale, psicoterapia cognitivo-comportamentale (CBT) in ambito comunitario

Tecnologia IA

Natural Language Processing (NLP) e Machine Learning applicati a sessioni di terapia

Algoritmo/Modello

Piattaforma Eleos Health (analisi trascrizioni e feedback su Evidence-Based Practices)

Compito/Utilizzo

Supporto decisionale e ottimizzazione (monitoraggio dei sintomi, feedback al terapeuta, generazione automatica di note cliniche)

Dataset

47 adulti con diagnosi di disturbi depressivi o d'ansia (23 in gruppo AI, 24 in TAU), dati real-world da una clinica comunitaria; trascrizioni di sedute e scale standardizzate (PHQ-9, GAD-7)

Outcome principali

Fattibilità e accettabilità della piattaforma; riduzione dei sintomi depressivi e ansiosi; aderenza alla terapia; qualità e tempistica delle note cliniche

Metri di valutazione

PHQ-9, GAD-7, numero medio di sedute completate, tempo medio di compilazione delle note cliniche, dimensione dell'effetto (Cohen's d)

Risultati chiave

I pazienti nel gruppo AI hanno partecipato al 67% di sedute in più rispetto a TAU. La riduzione dei sintomi depressivi è stata del 34% contro il 20% in TAU (d=0.82 vs 0.34). I sintomi ansiosi si sono ridotti del 29% contro l'8% in TAU (d=0.78 vs 0.14). I terapeuti hanno inviato le note cliniche 55 ore prima rispetto al gruppo TAU.

Benefici dichiarati

Migliore ritenzione dei pazienti, maggiore riduzione dei sintomi, supporto all'implementazione di pratiche basate su evidenze, riduzione del carico amministrativo per i terapeuti.

Limiti dichiarati

Campione ridotto (47 pazienti), durata breve (2 mesi), sintomi prevalentemente lievi-moderati, mancanza di follow-up a lungo termine, assenza di dati dettagliati sull'esperienza soggettiva dei pazienti.

Riferimento

Bibliografico

<https://doi.org/10.2196/46781>

Stratificazione per setting

AMBITO GENERALISTA AI IN SANITA'

Inclusione studio nella Revisione

NO

ID STUDIO:	ST40
TITOLO:	Wound imaging software and digital platform to assist review of surgical wounds using patient smartphones: The development and evaluation of artificial intelligence (WISDOM AI study)
Titolo in italiano:	Software di imaging delle ferite e piattaforma digitale per assistere la revisione delle ferite chirurgiche tramite smartphone dei pazienti: sviluppo e valutazione dell'intelligenza artificiale (studio WISDOM AI)
Autore/i	Melissa Rochon, Judith Tanner, James Jurkiewicz, Jacqueline Beckhelling, Akuha Aondoakaa, Keith Wilson, Luxmi Dhoonmoon, Max Underwood, Lara Mason, Roy Harris, Karen Cariaga
Anno	2024
Paese	Regno Unito
Tipo di studio	Studio prospettico di sviluppo e validazione
Contesto	Ambulatoriale e domiciliare, con follow-up post-chirurgico dei pazienti tramite smartphone
Specialità/Setting	Chirurgia generale e specialistica, controllo e monitoraggio delle ferite chirurgiche post-operatorie
Tecnologia IA	Computer vision basata su deep learning applicata all'imaging delle ferite
Algoritmo/Modello	Reti neurali convoluzionali sviluppate ad hoc per classificazione di immagini di ferite, integrate in una piattaforma digitale di telemedicina
Compito/Utilizzo	Supporto decisionale e monitoraggio remoto, con classificazione automatica delle immagini delle ferite in base al rischio di infezione o complicazioni post-operatorie
Dataset	Oltre 6.000 immagini di ferite chirurgiche raccolte da pazienti tramite smartphone in real-world setting clinico; le immagini sono state annotate da esperti per addestramento e validazione del modello
Outcome principali	Capacità del modello AI di identificare precocemente segni di infezione chirurgica e fornire supporto nella valutazione delle ferite a distanza; valutazione della fattibilità dell'uso di smartphone personali dei pazienti per garantire accessibilità e scalabilità del sistema
Metri di valutazione	Accuratezza complessiva superiore all'85%, sensibilità e specificità variabili in base alla tipologia di ferita, AUC riportata come superiore a 0,9 nei test di validazione; confronto con valutazioni cliniche standard
Risultati chiave	Il modello AI ha dimostrato performance elevate nella classificazione delle immagini e nella rilevazione precoce di complicanze, migliorando la possibilità di follow-up remoto e riducendo la necessità di visite in presenza; la piattaforma è stata ritenuta usabile dai pazienti e dagli operatori, mostrando potenziale per l'integrazione nella pratica clinica di routine
Benefici dichiarati	Monitoraggio tempestivo e accessibile delle ferite chirurgiche, riduzione delle complicanze grazie a diagnosi precoce, diminuzione della pressione sulle strutture ospedaliere attraverso follow-up remoto, miglioramento dell'aderenza dei pazienti al percorso di cura
Limiti dichiarati	Necessità di validazioni multicentriche e su popolazioni più ampie, possibili bias legati alla qualità delle immagini raccolte dai pazienti, rischio di disomogeneità nei dati dovuto alla varietà di smartphone utilizzati, bisogno di integrazione sicura nei sistemi clinici esistenti
Riferimento Bibliografico	https://doi.org/10.1371/journal.pone.0315384
Stratificazione per setting	AMBULATORIALE - CHIRURGICO
Inclusione studio nella Revisione	SI

ID STUDIO: ST41
TITOLO: Resource Allocation and Outpatient Appointment Scheduling Using Simulation Optimization

Titolo in italiano: Allocazione delle risorse e pianificazione degli appuntamenti ambulatoriali tramite ottimizzazione basata su simulazione

Autore/i	Carrie Ka Yuk Lin, Teresa Wai Ching Ling, Wing Kwan Yeung
Anno	2017
Paese	Hong Kong
Tipo di studio	Studio metodologico di simulazione con ottimizzazione
Contesto	Ambulatoriale
Specialità/Setting	Ambulatori generali e specialistici con gestione dei flussi di pazienti e risorse cliniche
Tecnologia IA	Ottimizzazione e modelli di simulazione integrati con tecniche di machine learning per la ricerca operativa
Algoritmo/Modello	Algoritmo di ottimizzazione basato su simulazione Monte Carlo e metaeuristiche applicate alla gestione degli appuntamenti; utilizzo di tecniche di analisi stocastica per valutare scenari alternativi
Compito/Utilizzo	Ottimizzazione della pianificazione degli appuntamenti e dell'allocazione delle risorse ambulatoriali per ridurre i tempi di attesa e migliorare l'efficienza operativa
Dataset	Dati real-world provenienti da un ambulatorio ospedaliero di Hong Kong, comprendenti flussi di pazienti, orari di appuntamenti e disponibilità di risorse; dataset di dimensione medio-grande con casi clinici raccolti in più mesi di attività
Outcome principali	Riduzione dei tempi di attesa dei pazienti, miglior utilizzo delle risorse ambulatoriali, incremento della produttività complessiva del sistema; valutazione dell'efficacia delle strategie di scheduling ottimizzate rispetto ai metodi tradizionali
Metri di valutazione	Tempi medi di attesa dei pazienti, tasso di utilizzo delle risorse, numero di appuntamenti completati, miglioramento percentuale rispetto agli scenari baseline
Risultati chiave	L'uso di simulazione e ottimizzazione ha permesso di ottenere riduzioni significative dei tempi di attesa, un aumento nell'utilizzo delle risorse cliniche e una gestione più bilanciata dei flussi di pazienti; i modelli proposti hanno dimostrato robustezza in scenari multipli, suggerendo la possibilità di implementazione reale per migliorare l'efficienza dei servizi ambulatoriali
Benefici dichiarati	Maggior efficienza operativa, riduzione dei tempi di attesa, incremento della soddisfazione dei pazienti, migliore distribuzione delle risorse umane e tecnologiche
Limiti dichiarati	Applicabilità limitata a specifici contesti ospedalieri, necessità di validazioni prospettiche in strutture diverse, rischio di ridotta generalizzabilità dovuto alle particolarità organizzative dell'ospedale analizzato
Riferimento Bibliografico	https://doi.org/10.1155/2017/9034737
Stratificazione per setting	AMBULATORIALE
Inclusione studio nella Revisione	SI

ID STUDIO: ST42
TITOLO: Machine learning-augmented interventions in perioperative care: a systematic review and meta-analysis

Titolo in italiano: Interventi potenziati dal machine learning nella cura perioperatoria: revisione sistematica e meta-analisi

Autore/i	Divya Mehta, Xiomara T. Gonzalez, Grace Huang, Joanna Abraham
Anno	2024
Paese	Stati Uniti
Tipo di studio	Revisione sistematica con meta-analisi
Contesto	Chirurgico, con applicazioni in ambito perioperatorio
Specialità/Setting	Chirurgia generale e specialistica, anesthesiologia, medicina perioperatoria
Tecnologia IA	Machine learning applicato a modelli predittivi e sistemi di supporto decisionale in chirurgia
Algoritmo/Modello	Random forest, support vector machine, gradient boosting, reti neurali profonde, modelli ensemble; vari algoritmi confrontati nei diversi studi inclusi
Compito/Utilizzo	Previsione di esiti perioperatori come complicanze, mortalità, riammissione e durata della degenza; supporto decisionale per la gestione clinica dei pazienti chirurgici
Dataset	Dati clinici real-world provenienti da cartelle elettroniche, database ospedalieri e registri chirurgici; dimensione dei dataset variabile da poche centinaia a decine di migliaia di pazienti, con prevalenza di studi retrospettivi
Outcome principali	Capacità predittiva dei modelli di machine learning sugli esiti chirurgici; confronto delle performance rispetto ai modelli statistici tradizionali; valutazione del potenziale clinico degli algoritmi
Metri di valutazione	Area under the curve (AUC), accuratezza, sensibilità, specificità, odds ratio e hazard ratio nei modelli meta-analitici
Risultati chiave	I modelli di machine learning hanno mostrato performance predittive superiori ai metodi statistici convenzionali, con AUC mediamente più elevati nella previsione di complicanze e riammissioni post-operatorie; le evidenze suggeriscono un potenziale clinico rilevante ma sottolineano la scarsità di validazioni prospettiche e di implementazioni reali nei flussi chirurgici
Benefici dichiarati	Migliore capacità di stratificare il rischio perioperatorio, supporto alle decisioni cliniche per anestesisti e chirurghi, possibilità di personalizzare la gestione dei pazienti chirurgici, riduzione di complicanze e costi ospedalieri
Limiti dichiarati	Predominanza di studi retrospettivi, eterogeneità degli algoritmi e degli outcome considerati, limitata disponibilità di validazioni esterne e studi prospettici, scarsa generalizzabilità a diversi contesti clinici
Riferimento Bibliografico	https://doi.org/10.1016/j.bja.2024.08.007
Stratificazione per setting	CHIRURGICO
Inclusione studio nella Revisione	SI

ID STUDIO: ST43
TITOLO: Artificial intelligence for surgical services in Australia and New Zealand: opportunities, challenges and recommendation

Titolo in italiano: Intelligenza artificiale nei servizi chirurgici in Australia e Nuova Zelanda:

opportunità, sfide e raccomandazioni

Autore/i	Joshua G. Kooor, Stephen Bacchi, Prakriti Sharma, Srishti Sharma, Medhir Kumawat, Brandon Stretton, Aashray K. Gupta, WengOnn Chan, Amal Abou-Hamden, Guy J. Maddern
Anno	2024
Paese	Australia e Nuova Zelanda
Tipo di studio	Articolo di prospettiva / analisi critica con raccomandazioni
Contesto	Chirurgico
Specialità/Setting	Chirurgia generale e specialistica, con applicazioni perioperatorie e a livello di sistema sanitario
Tecnologia IA	Machine learning, deep learning e strumenti di computer vision, con applicazioni predittive e di supporto decisionale
Algoritmo/Modello	Diversi modelli AI richiamati (Adelaide Score, strumenti predittivi e di visione intraoperatoria), senza un singolo algoritmo sviluppato nello studio ma con enfasi su implementazioni già presenti e in sviluppo
Compito/Utilizzo	Supporto decisionale, previsione di esiti, ottimizzazione del percorso chirurgico e monitoraggio post-operatorio
Dataset	Evidenze prevalentemente da studi preliminari e validazioni locali (es. Adelaide Score); si sottolinea la necessità di dataset real-world e di studi prospettici su popolazioni australiane e neozelandesi
Outcome principali	Discussione sulle potenzialità dell'IA nel migliorare efficienza ed esiti dei pazienti chirurgici, evidenziando però la scarsità di trial prospettici e la necessità di validazioni locali; identificazione dei rischi etici e organizzativi legati alla sua applicazione
Metri di valutazione	Non applicabili in quanto articolo prospettico; vengono richiamate AUC e metriche predittive di altri studi già pubblicati, ma non prodotte nuove valutazioni quantitative
Risultati chiave	L'IA ha il potenziale di migliorare la precisione intraoperatoria, l'accuratezza diagnostica e la gestione del rischio perioperatorio, ma il suo impiego richiede regolamentazione, monitoraggio post-implementazione, audit continuo e forte governance etica; l'articolo fornisce raccomandazioni operative per un'adozione sicura
Benefici dichiarati	Maggior efficienza, miglioramento degli esiti clinici, possibilità di ridurre disuguaglianze geografiche nell'accesso ai servizi, potenziamento del supporto decisionale per i chirurghi
Limiti dichiarati	Mancanza di trial randomizzati e validazioni real-world, rischio di bias nei dati di addestramento, scarsa trasparenza degli algoritmi (black box), possibili problemi etici e di responsabilità medico-legale
Riferimento Bibliografico	https://doi.org/10.5694/mja2.52225
Stratificazione per setting	CHIRURGICO
Inclusione studio nella Revisione	SI

ID STUDIO: ST44

TITOLO: Enhancing Surgery Scheduling in Health Care Settings With Metaheuristic Optimization Models: Algorithm Validation Study

Titolo in italiano: Migliorare la pianificazione chirurgica nelle strutture sanitarie con modelli di ottimizzazione metaeuristica: studio di validazione

algoritmica

Autore/i	João Lopes, Tiago Guimarães, Júlio Duarte, Manuel Santos
Anno	2025
Paese	Portogallo
Tipo di studio	Studio di validazione algoritmica
Contesto	Chirurgico
Specialità/Setting	Pianificazione delle liste operatorie in ospedali con risorse limitate e complessità organizzativa
Tecnologia IA	Ottimizzazione metaeuristica applicata alla pianificazione chirurgica
Algoritmo/Modello	Modelli di metaeuristica ibridi basati su algoritmi genetici, simulated annealing e tabu search, combinati per affrontare il problema del job-shop scheduling chirurgico
Compito/Utilizzo	Ottimizzazione della programmazione chirurgica per ridurre tempi di attesa, aumentare l'utilizzo delle sale operatorie e migliorare l'efficienza complessiva del sistema
Dataset	Dati real-world di un ospedale portoghese comprendenti programmazioni chirurgiche, durata media degli interventi, disponibilità di sale e vincoli organizzativi; dataset di più anni utilizzato per addestramento e validazione del modello
Outcome principali	Dimostrazione che i modelli metaeuristici migliorano la performance di scheduling rispetto agli approcci tradizionali, garantendo maggiore equilibrio tra efficienza operativa e qualità del servizio ai pazienti
Metri di valutazione	Tempi medi di attesa, tasso di utilizzo delle sale operatorie, numero di interventi completati, confronto percentuale con metodi di scheduling convenzionali
Risultati chiave	L'approccio basato su metaeuristiche ha ridotto significativamente i tempi di attesa dei pazienti, migliorato l'uso delle sale operatorie e ottimizzato la distribuzione degli interventi rispetto ai metodi manuali o tradizionali, dimostrando robustezza e adattabilità a scenari complessi
Benefici dichiarati	Maggior efficienza nella gestione delle risorse chirurgiche, riduzione dei tempi di lista d'attesa, aumento della capacità operativa degli ospedali, supporto alle decisioni per i manager sanitari
Limiti dichiarati	Necessità di ulteriori validazioni in contesti diversi, rischio di ridotta generalizzabilità dovuto alle specificità organizzative dell'ospedale utilizzato per la validazione, complessità computazionale dei modelli in scenari molto ampi
Riferimento Bibliografico	https://doi.org/10.2196/57231
Stratificazione per setting	CHIRURGICO
Inclusione studio nella Revisione	SI

ID STUDIO:	ST45
TITOLO:	Improving Operating Room Efficiency: A machine learning approach to predict case-time duration
Titolo in italiano:	Migliorare l'efficienza della sala operatoria: un approccio di machine learning per predire la durata degli interventi chirurgici
Autore/i	Matthew A. Bartek, Rajeev C. Saxena, Stuart Solomon, Christine T. Fong, Lakshmana D. Behara, Ravitheja Venigandla, Kalyani Velagapudi, Bala G. Nair, John D. Lang

Anno	2019
Paese	Stati Uniti
Tipo di studio	Studio osservazionale retrospettivo con sviluppo e validazione di modelli predittivi
Contesto	Chirurgico
Specialità/Setting	Sala operatoria in ospedale terziario universitario (University of Washington Medical Center)
Tecnologia IA	Machine Learning (modelli supervisionati)
Algoritmo/Modello	Random Forest e XGBoost (Extreme Gradient Boosting), confrontati con regressione lineare e stime convenzionali Compito: Previsione della durata degli interventi chirurgici per migliorare la programmazione operatoria
Compito/Utilizzo	Previsione della durata degli interventi chirurgici per migliorare la programmazione operatoria
Dataset	46.986 interventi chirurgici programmati tra il 2014 e il 2017; suddivisi in training set (80%) e test set (20%); dati real-world estratti dal sistema informativo perioperatorio e cartella clinica elettronica
Outcome principali	Accuratezza delle previsioni rispetto agli standard istituzionali; riduzione di sovrastime (overage) e sottostime (underage); percentuale di casi con durata effettiva entro $\pm 10\%$ della stima
Metri di valutazione	R^2 , MAPE (Mean Absolute Percentage Error), percentuale di casi entro $\pm 10\%$ di accuratezza, percentuali di overage e underage
Risultati chiave	I modelli ML hanno superato nettamente le stime dei chirurghi e quelle basate sulla media storica. Il modello XGBoost specifico per chirurgo ha migliorato la capacità predittiva, con il 39% di casi entro la soglia $\pm 10\%$ (vs 32% dei chirurghi). Alcuni modelli chirurgici individuali hanno raggiunto fino al 50% di accuratezza entro la soglia. Le variabili più predittive erano legate a procedure e chirurghi (durata media di interventi simili), mentre i fattori clinici del paziente avevano un peso minore.
Benefici dichiarati	Miglioramento dell'efficienza delle sale operatorie, riduzione di ritardi e straordinari, migliore pianificazione delle risorse e potenziale riduzione dei costi sanitari.
Limiti dichiarati	Studio monocentrico in ospedale ad alta complessità; generalizzabilità limitata; variabilità tra chirurghi; il miglioramento predittivo non implica automaticamente un aumento dell'efficienza globale delle sale operatorie; non valutato l'impatto economico diretto.
Riferimento Bibliografico	https://doi.org/10.1016/j.jamcollsurg.2019.05.029
Stratificazione per setting	CHIRURGICO
Inclusione studio nella Revisione	SI

ID STUDIO:	ST46
TITOLO:	Leveraging machine learning to enhance appointment adherence at a novel post-discharge care transition clinic
Titolo in italiano:	Utilizzo del machine learning per migliorare l'aderenza agli appuntamenti in una nuova clinica di transizione post-dimissione
Autore/i	Seung-Yup Lee, Reid M. Eagleson, Larry R. Hearld, Madeline J. Gibson, Kristine R. Hearld, Allyson G. Hall, Greer A. Burkholder, Jacob McMahon, Shoab Y.

	Mahmood, Corey T. Spraberry, Thalia J. Baker, Alison R. Garretson, Heather M. Bradley, Michael J. Mugavero
Anno	2024
Paese	Stati Uniti
Tipo di studio	Studio osservazionale retrospettivo con sviluppo di modello predittivo
Contesto	Ambulatoriale, clinica di transizione post-dimissione ospedaliera
Specialità/Setting	Medicina interna, gestione della continuità assistenziale e follow-up dopo dimissione ospedaliera
Tecnologia IA	Machine learning supervisionato
Algoritmo/Modello	Gradient Boosting Machine (GBM), confrontato con regressione logistica e altri modelli tradizionali
Compito/Utilizzo	Previsione dell'aderenza agli appuntamenti post-dimissione per identificare pazienti a rischio di mancata presentazione e migliorare la gestione delle risorse cliniche
Dataset	1.170 pazienti dimessi e indirizzati a una clinica di transizione post-acuta; dati real-world estratti da cartelle cliniche elettroniche e registri ambulatoriali; set di training e test separati per valutazione del modello
Outcome principali	Accuratezza predittiva del modello ML sull'aderenza agli appuntamenti; identificazione delle variabili più rilevanti per la previsione (età, condizioni cliniche croniche, caratteristiche socio-demografiche, storia di aderenza precedente)
Metri di valutazione	Area under the curve (AUC), sensibilità, specificità, importanza delle variabili nei modelli predittivi
Risultati chiave	Il modello Gradient Boosting ha ottenuto un'AUC pari a 0,77, superiore rispetto alla regressione logistica e ad altri algoritmi testati. Le variabili più predittive risultavano essere l'età, la comorbidità e la storia di aderenza agli appuntamenti. L'uso del modello ha mostrato potenziale per stratificare i pazienti a rischio e implementare interventi mirati di promemoria o supporto.
Benefici dichiarati	Possibilità di migliorare l'aderenza agli appuntamenti post-dimissione, ridurre riospedalizzazioni e costi sanitari, supportare una migliore continuità assistenziale e ottimizzare l'uso delle risorse cliniche.
Limiti dichiarati	Studio monocentrico con campione relativamente ridotto, necessità di validazioni esterne in altri contesti clinici, rischio di bias legato alle variabili socio-demografiche, assenza di valutazione prospettica dell'impatto del modello sulla pratica reale
Riferimento Bibliografico	https://doi.org/10.1093/jamiaopen/ooae086
Stratificazione per setting	AMBULATORIALE
Inclusione studio nella Revisione	SI

ID STUDIO: ST47

TITOLO: Artificial Intelligence and its role in surgical care in low-income and middle-income countries

Titolo in italiano: L'intelligenza artificiale e il suo ruolo nell'assistenza chirurgica nei paesi a basso e medio reddito

Autore/i	Ché L. Reddy, Shivani Mitra, John G. Meara, Rifat Atun, Salim Afshar
Anno	2019
Paese	Stati Uniti (con focus applicativo su paesi a basso e medio reddito: Brasile, Cina,

Tipo di studio	India, Sudafrica, Turchia)
Contesto	Commento/analisi critica in rivista peer-reviewed (The Lancet Digital Health)
Specialità/Setting	Chirurgico, con attenzione ai sistemi sanitari dei paesi a risorse limitate
Tecnologia IA	Chirurgia generale e specialistica in paesi LMIC, con enfasi sui sistemi sanitari piuttosto che sulla singola pratica clinica
Algoritmo/Modello	Machine learning e deep learning applicati a dati sanitari e sistemi gestionali
Compito/Utilizzo	Nessun algoritmo sviluppato, discussione concettuale su applicazioni predittive, diagnostiche e di supporto decisionale attraverso AI trasparente e spiegabile
Dataset	Ottimizzazione dei sistemi chirurgici, supporto decisionale e rafforzamento dei processi organizzativi (supply chain, workforce, governance) secondo il framework National Surgical, Obstetric, and Anesthesia Planning (NSOAP)
Outcome principali	Non presenta un dataset originale, ma discute la possibilità di usare grandi repository di dati già presenti in paesi come Brasile, India e Sudafrica; enfasi sulla gestione di dati "asimmetrici, asincroni e distribuiti" piuttosto che su dataset centralizzati
Metri di valutazione	Identificazione delle potenzialità dell'IA per affrontare barriere sistemiche alla chirurgia nei LMIC, in particolare carenze infrastrutturali, di personale e di governance; raccomandazioni per usare AI a supporto della pianificazione sanitaria e dell'efficienza dei processi
Risultati chiave	Non applicabili (assenza di sperimentazione quantitativa)
Benefici dichiarati	L'IA, se implementata in modo trasparente e adattato al contesto, potrebbe rafforzare i sistemi chirurgici LMIC migliorando gestione delle risorse, allocazione del personale e pianificazione dei servizi; il potenziale è significativo ma dipende dalla qualità e governance dei dati, nonché da capacità locali di sviluppo tecnologico
Limiti dichiarati	Potenziale di ridurre inefficienze burocratiche, ottimizzare supply chain e workforce, migliorare equità nell'accesso chirurgico; possibilità di sfruttare dataset già esistenti in modo più intelligente
Riferimento Bibliografico	Mancanza di studi empirici, problemi di qualità e standardizzazione dei dati, rischio di trasferimento passivo di strumenti sviluppati in HIC (high-income countries) non adattati al contesto LMIC, questioni etiche e regolatorie ancora irrisolte
Stratificazione per setting	https://doi.org/10.1016/S2589-7500(19)30200-6
Inclusione studio nella Revisione	CHIRURGICO
	NO

ID STUDIO:	ST48
TITOLO:	Transforming Healthcare in Low-Resource Settings With Artificial Intelligence: Recent Developments and Outcomes
Titolo in italiano:	Trasformare l'assistenza sanitaria nei contesti a basse risorse con l'intelligenza artificiale: sviluppi recenti e risultati
Autore/i	Ravi Rai Dangi, Anil Sharma, Vipin Vageriya
Anno	2024
Paese	India
Tipo di studio	Revisione narrativa con casi applicativi
Contesto	Ambulatoriale e ospedaliero, con particolare attenzione a contesti rurali e a basso reddito

Specialità/Setting	Medicina generale, telemedicina, diagnostica clinica e radiologica, salute materno-infantile, malattie infettive
Tecnologia IA	Machine learning, deep learning, natural language processing e computer vision applicati a diagnosi e monitoraggio
Algoritmo/Modello	Applicazioni basate su reti neurali convoluzionali per imaging medico, NLP per triage e chatbot sanitari, modelli predittivi di machine learning per diagnosi precoce e stratificazione del rischio; non viene sviluppato un singolo algoritmo originale ma riportate diverse implementazioni in ambito LMIC
Compito/Utilizzo	Diagnosi precoce, supporto decisionale, telemonitoraggio, triage automatico, ottimizzazione delle risorse e formazione del personale sanitario tramite strumenti AI-based
Dataset	Evidenze raccolte da programmi pilota e iniziative real-world in India e altri paesi a basso reddito; dataset eterogenei di immagini mediche, referti clinici e dati epidemiologici; molti studi citati riportano casistiche da centinaia a migliaia di pazienti
Outcome principali	L'IA ha migliorato l'accessibilità ai servizi diagnostici, la tempestività delle cure e la gestione delle malattie croniche e infettive in contesti rurali; incremento dell'accuratezza diagnostica, riduzione dei tempi di attesa e miglioramento dell'efficienza operativa
Metri di valutazione	Accuratezza diagnostica, sensibilità, specificità, AUC; in alcuni casi miglioramento percentuale dell'aderenza alle cure e della tempestività degli interventi rispetto agli standard tradizionali
Risultati chiave	L'adozione di AI ha permesso diagnosi più rapide e affidabili in assenza di specialisti, ha supportato la telemedicina riducendo le distanze tra pazienti e strutture sanitarie e ha migliorato la gestione dei programmi di salute pubblica; si evidenzia tuttavia la mancanza di validazioni multicentriche e la necessità di maggiore infrastruttura digitale
Benefici dichiarati	Aumento dell'accesso ai servizi sanitari, riduzione delle disuguaglianze geografiche, supporto agli operatori sanitari locali, miglioramento degli esiti clinici attraverso diagnosi precoce e personalizzata
Limiti dichiarati	Barriere infrastrutturali, scarsità di dataset di qualità, rischio di bias culturali e linguistici nei modelli NLP, mancanza di standardizzazione e validazioni robuste, problemi di sostenibilità economica
Riferimento Bibliografico	https://doi.org/10.1111/phn.13500
Stratificazione per setting	AMBITO GENERALISTA AI IN SANITA
Inclusione studio nella Revisione	NO

ID STUDIO:	ST49
TITOLO:	The role of artificial intelligence in surgical patient perioperative management
Titolo in italiano:	Il ruolo dell'intelligenza artificiale nella gestione perioperatoria del paziente chirurgico
Autore/i	Elena G. Bignami, Federico Cozzani, Paolo Del Rio, Valentina Bellini
Anno	2021
Paese	Italia
Tipo di studio	Expert opinion/revisione narrativa (Minerva Anestesiologica)
Contesto	Chirurgico, gestione perioperatoria

Specialità/Setting	Anestesia, chirurgia generale e perioperatoria in ospedale universitario
Tecnologia IA	Machine learning e deep learning applicati a EHR, imaging, monitoraggio perioperatorio e analisi video intraoperatori
Algoritmo/Modello	Diversi esempi riportati dalla letteratura: MySurgeryRisk per predire complicanze postoperatorie, modelli ML per stratificazione rischio, Hypotension Prediction Index per ipotensione intraoperatoria, reti neurali convoluzionali per imaging intraoperatorio
Compito/Utilizzo	Predizione di complicanze postoperatorie, supporto decisionale in fase preoperatoria, monitoraggio intraoperatorio, diagnosi intraoperatoria, previsione di eventi avversi e supporto alla dimissione sicura
Dataset	Studi citati riportano grandi coorti retrospettive (es. 51.457 pazienti, 66.370 pazienti, oltre 500.000 minuti di monitoraggio emodinamico); dati real-world provenienti da cartelle cliniche elettroniche e sistemi di monitoraggio intraoperatorio
Outcome principali	Miglioramento della capacità predittiva per complicanze, ipotensione, ipossiemia e riammissioni; maggiore personalizzazione nella gestione del paziente chirurgico rispetto ai calcolatori di rischio tradizionali
Metri di valutazione	AUC (0,85–0,87 per predizione riammissione), accuratezza predittiva, confronto con ACS NSQIP e altri calcolatori tradizionali; superiorità dei modelli ML rispetto a metodi convenzionali
Risultati chiave	L'intelligenza artificiale può rafforzare la medicina perioperatoria in tutte le fasi (pre, intra e post-operatoria), migliorando la precisione predittiva e consentendo un approccio più personalizzato. Sono già disponibili modelli con performance superiori agli strumenti tradizionali, specialmente per rischio di complicanze e gestione anestesiológica predittiva.
Benefici dichiarati	Ottimizzazione della gestione perioperatoria, riduzione complicanze, miglioramento della sicurezza del paziente, sostegno alla standardizzazione dei percorsi perioperatori e potenziale riduzione dei costi sanitari.
Limiti dichiarati	Necessità di dataset di alta qualità e adeguata numerosità, difficoltà di standardizzazione, complessità di integrazione nella pratica clinica quotidiana, mancanza di formazione specifica degli operatori, problematiche etiche e legislative ancora aperte.
Riferimento Bibliografico	https://doi.org/10.23736/S0375-9393.20.14999-X
Stratificazione per setting	CHIRURGICO
Inclusione studio nella Revisione	SI

ID STUDIO: ST50

TITOLO: Artificial Intelligence and Robotics in Spine Surgery

Titolo in italiano: Intelligenza artificiale e robotica nella chirurgia spinale

Autore/i	Jonathan J. Rasouli, Jianning Shao, Sean Neifert, Wende N. Gibbs, Ghaith Habboub, Michael P. Steinmetz, Edward Benzal, Thomas E. Mroz
Anno	2021
Paese	Stati Uniti
Tipo di studio	Revisione narrativa
Contesto	Chirurgico
Specialità/Setting	Neurochirurgia e chirurgia spinale
Tecnologia IA	Machine learning, deep learning, computer vision, robotica chirurgica assistita da

Algoritmo/Modello	AI Applicazioni di reti neurali convoluzionali (CNN) per analisi di imaging, algoritmi di ML per predizione di complicanze, sistemi robotici assistiti da AI per posizionamento di viti peduncolari e pianificazione chirurgica
Compito/Utilizzo	Supporto decisionale pre-operatorio, analisi di imaging, pianificazione chirurgica, ottimizzazione intraoperatoria, esecuzione robot-assistita di procedure spinali
Dataset	Studi citati comprendono dataset di migliaia di immagini radiologiche e dati clinici retrospettivi; focus su dati real-world per addestramento di modelli di imaging e validazioni di sistemi robotici
Outcome principali	Maggiore accuratezza diagnostica nell'analisi radiologica, precisione nel posizionamento delle viti peduncolari, riduzione delle complicanze e miglioramento della sicurezza intraoperatoria; possibilità di personalizzare la chirurgia in base al rischio predetto
Metri di valutazione	Accuratezza predittiva, AUC per modelli diagnostici, tassi di successo e precisione del posizionamento delle viti, confronto tra chirurgia tradizionale e robot-assistita
Risultati chiave	L'integrazione di AI e robotica nella chirurgia spinale migliora la precisione tecnica e la sicurezza del paziente; i modelli predittivi basati su AI superano gli score clinici tradizionali nella stratificazione del rischio e le tecnologie robotiche mostrano tassi di accuratezza superiori rispetto al freehand tradizionale
Benefici dichiarati	Miglioramento degli outcome chirurgici, riduzione del margine di errore intraoperatorio, standardizzazione delle procedure, supporto al training dei chirurghi, potenziale riduzione dei costi a lungo termine grazie a minori complicanze
Limiti dichiarati	Elevati costi iniziali di implementazione, necessità di dataset di grandi dimensioni e di alta qualità per l'addestramento, mancanza di studi prospettici randomizzati, problematiche etiche e medico-legali legate all'automazione chirurgica
Riferimento Bibliografico	https://doi.org/10.1177/2192568220915718
Stratificazione per setting	CHIRURGICO
Inclusione studio nella Revisione	SI

ID STUDIO:	ST51
TITOLO:	Machine learning based integrated scheduling and rescheduling for elective and emergency patients in the operating theatre
Titolo in italiano:	Pianificazione e ripianificazione integrata basata su machine learning per pazienti elettivi ed emergenti in sala operatoria
Autore/i	Masoud Eshghali, Devika Kannan, Navid Salmanzadeh-Meydani, Amir Mohammad Esmayeeli Sikaroudi
Anno	2024
Paese	Danimarca e Iran (collaborazione internazionale)
Tipo di studio	Studio metodologico con sviluppo e validazione di modello ML per ottimizzazione organizzativa
Contesto	Chirurgico
Specialità/Setting	Gestione delle liste operatorie per pazienti elettivi ed emergenti in ospedali pubblici
Tecnologia IA	Machine learning integrato con tecniche di ottimizzazione matematica
Algoritmo/Modello	Modelli ML per predire la durata degli interventi combinati con Mixed-Integer

Compito/Utilizzo	Linear Programming (MILP) e tecniche di ottimizzazione stocastica per scheduling dinamico; integrazione di gradient boosting per previsioni accurate dei tempi chirurgici
Dataset	Ottimizzazione della pianificazione chirurgica e della ripianificazione dinamica in presenza di urgenze, con l'obiettivo di minimizzare ritardi, cancellazioni e tempi morti delle sale operatorie
Outcome principali	Dati real-world di interventi chirurgici elettivi ed emergenti provenienti da ospedali universitari; migliaia di casi analizzati con caratteristiche cliniche e operative; dataset suddiviso in training e test per la validazione del modello
Metri di valutazione	Riduzione dei ritardi nelle liste operatorie, miglioramento dell'uso delle sale, diminuzione delle cancellazioni e incremento della flessibilità organizzativa; confronto dei risultati con modelli di scheduling tradizionali
Risultati chiave	Percentuale di riduzione dei ritardi, tasso di utilizzo delle sale, accuratezza delle previsioni sui tempi chirurgici (MAPE, RMSE), confronto tra scenari simulati e baseline reale
Benefici dichiarati	L'approccio ML integrato ha superato i metodi tradizionali di scheduling, mostrando riduzioni significative dei ritardi e un migliore adattamento agli imprevisti legati agli ingressi di emergenza. I modelli hanno garantito una maggiore stabilità organizzativa e un utilizzo più efficiente delle risorse ospedaliere.
Limiti dichiarati	Ottimizzazione della gestione delle risorse chirurgiche, riduzione delle inefficienze organizzative, incremento della capacità operativa, supporto alle decisioni manageriali in tempo reale
Riferimento Bibliografico	Complessità computazionale elevata, necessità di validazioni multicentriche, rischio di ridotta generalizzabilità in contesti con diversa organizzazione sanitaria, dipendenza dalla qualità dei dati inseriti
Stratificazione per setting	https://doi.org/10.1007/s10479-023-05168-x
Inclusione studio nella Revisione	CHIRURGICO
	SI

ID STUDIO:	ST52
TITOLO:	Artificial Intelligence: A New Tool in Operating Room Management. Role of Machine Learning Models in Operating Room Optimization
Titolo in italiano:	Intelligenza artificiale: un nuovo strumento nella gestione della sala operatoria. Ruolo dei modelli di machine learning nell'ottimizzazione del blocco operatorio
Autore/i	Valentina Bellini, Marco Guzzon, Barbara Bigliardi, Monica Mordonini, Serena Filippelli, Elena Bignami
Anno	2020
Paese	Italia
Tipo di studio	Revisione sistematica con prospettive applicative
Contesto	Chirurgico
Specialità/Setting	Gestione del blocco operatorio e ottimizzazione perioperatoria in ospedale
Tecnologia IA	Machine learning applicato a sistemi predittivi e di supporto decisionale
Algoritmo/Modello	Vengono discussi diversi modelli ML (support vector machine, random forest, reti neurali) applicati a predizione di tempi chirurgici, complicanze perioperatorie e gestione risorse; nessun algoritmo originale sviluppato nello studio

Compito/Utilizzo	Ottimizzazione della gestione del blocco operatorio tramite predizione della durata degli interventi, analisi dei fattori di rischio e supporto alla pianificazione delle risorse
Dataset	Studi citati utilizzano dataset retrospettivi di ospedali con migliaia di interventi; dati real-world estratti da EHR e sistemi informativi perioperatori; non viene utilizzato un dataset originale ma riportati risultati della letteratura
Outcome principali	L'IA mostra potenziale nell'aumentare accuratezza delle previsioni sui tempi operatori, nel ridurre ritardi e inefficienze organizzative e nel migliorare la gestione dei rischi perioperatori
Metri di valutazione	Accuratezza predittiva, AUC, MAPE e confronto con stime tradizionali; evidenziata la superiorità dei modelli ML rispetto ai sistemi convenzionali
Risultati chiave	L'intelligenza artificiale è vista come strumento capace di trasformare la gestione della sala operatoria, migliorando la capacità di pianificazione e di adattamento in tempo reale alle variabili cliniche e organizzative. Viene sottolineato il potenziale dei modelli predittivi di ML per anticipare criticità e supportare decisioni manageriali.
Benefici dichiarati	Miglioramento dell'efficienza e della produttività delle sale operatorie, riduzione dei ritardi e dei costi, ottimizzazione della sicurezza del paziente e del lavoro degli operatori sanitari
Limiti dichiarati	Necessità di dataset ampi e di alta qualità, mancanza di validazioni prospettiche, difficoltà di integrazione dei modelli ML nei sistemi clinici esistenti, questioni etiche e di accettazione da parte del personale
Riferimento Bibliografico	https://doi.org/10.1007/s10916-019-1512-1
Stratificazione per setting	CHIRURGICO
Inclusione studio nella Revisione	SI

ID STUDIO: ST53

TITOLO: Artificial Intelligence in Operating Room Management

Titolo in italiano: Intelligenza artificiale nella gestione della sala operatoria

Autore/i	Valentina Bellini, Michele Russo, Tania Domenichetti, Matteo Panizzi, Simone Allai, Elena Giovanna Bignami
Anno	2024
Paese	Italia
Tipo di studio	Revisione sistematica con prospettive applicative
Contesto	Chirurgico
Specialità/Setting	Gestione del blocco operatorio in ospedali universitari e di alta complessità
Tecnologia IA	Machine learning, deep learning, natural language processing, computer vision applicati a modelli predittivi e gestionali
Algoritmo/Modello	Discussione di modelli ML supervisionati (random forest, gradient boosting, reti neurali), NLP per l'analisi di referti clinici e CV per l'analisi video intraoperatoria; non è sviluppato un algoritmo originale ma vengono riportati esempi applicativi dalla letteratura
Compito/Utilizzo	Previsione dei tempi operatori, supporto decisionale perioperatorio, gestione dinamica delle risorse, rilevazione automatica di eventi intraoperatori, ottimizzazione del percorso chirurgico
Dataset	Evidenze tratte da studi precedenti su dataset real-world di cartelle cliniche elettroniche, registri chirurgici e sistemi di monitoraggio; casistiche variabili da

Outcome principali	migliaia a decine di migliaia di interventi, con dati retrospettivi e prospettici L'IA può migliorare la precisione delle previsioni sui tempi operatori e sugli esiti perioperatori, ottimizzare l'uso delle risorse e ridurre inefficienze e ritardi; il potenziale clinico-organizzativo è alto ma dipende dall'integrazione nei flussi reali
Metri di valutazione	AUC, accuratezza, sensibilità, specificità, MAPE, confronto con stime tradizionali; riportati valori mediamente superiori dei modelli ML rispetto ai sistemi convenzionali
Risultati chiave	L'IA è in grado di supportare una gestione più efficiente e predittiva del blocco operatorio, integrando diversi approcci tecnologici (ML, NLP, CV) e facilitando decisioni in tempo reale; tuttavia permangono barriere legate a standardizzazione, accettazione clinica ed etica
Benefici dichiarati	Miglioramento della pianificazione chirurgica, riduzione dei ritardi, ottimizzazione dei costi e delle risorse, maggiore sicurezza del paziente, riduzione del carico organizzativo per il personale
Limiti dichiarati	Mancanza di validazioni prospettiche multicentriche, difficoltà di implementazione nei sistemi ospedalieri esistenti, necessità di grandi dataset standardizzati e di formazione specifica del personale, questioni etiche e di responsabilità
Riferimento Bibliografico	https://doi.org/10.1007/s10916-024-02038-2
Stratificazione per setting	CHIRURGICO
Inclusione studio nella Revisione	SI

ID STUDIO:	ST54
TITOLO:	A comprehensive review on operating room scheduling and optimization
Titolo in italiano:	Una revisione completa sulla pianificazione e l'ottimizzazione della sala operatoria

Autore/i	Md Al Amin, Roberto Baldacci, Vahid Kayvanfar
Anno	2024
Paese	Francia e Iran (collaborazione internazionale)
Tipo di studio	Revisione sistematica della letteratura
Contesto	Chirurgico
Specialità/Setting	Gestione della sala operatoria e ottimizzazione dei processi ospedalieri
Tecnologia IA	Machine learning, deep learning e modelli di ottimizzazione matematica applicati alla programmazione operatoria
Algoritmo/Modello	La revisione include algoritmi ML supervisionati e non supervisionati (random forest, SVM, reti neurali), modelli di ottimizzazione (Mixed-Integer Linear Programming, metaeuristiche come simulated annealing e genetic algorithms) e approcci ibridi ML-ottimizzazione
Compito/Utilizzo	Analisi critica degli approcci di scheduling e ottimizzazione per ridurre ritardi, aumentare efficienza e migliorare l'utilizzo delle risorse operatorie
Dataset	Studi analizzati provenienti da casistiche real-world ospedaliere e da scenari simulati; dimensioni variabili da centinaia a decine di migliaia di interventi; focus sulla distinzione tra approcci basati su dati reali e modelli teorici
Outcome principali	Identificazione delle tendenze emergenti nell'uso dell'IA per la pianificazione chirurgica, con enfasi sugli approcci ibridi ML-ottimizzazione e sulla necessità di

Metri di valutazione	validazioni prospettiche; la revisione mostra che i metodi AI superano costantemente i sistemi tradizionali in termini di accuratezza e adattabilità Accuratezza predittiva, MAPE, RMSE, AUC, tempi di utilizzo della sala operatoria, percentuali di riduzione dei ritardi; metriche dipendenti dai singoli studi analizzati
Risultati chiave	L'integrazione di AI e modelli di ottimizzazione ha migliorato sensibilmente la gestione delle sale operatorie, riducendo tempi morti, cancellazioni e ritardi. La revisione sottolinea la progressiva evoluzione dagli approcci tradizionali a quelli ibridi, che combinano predizione ML e algoritmi matematici di scheduling dinamico
Benefici dichiarati	Maggior efficienza organizzativa, miglioramento dell'esperienza del paziente, ottimizzazione dell'uso delle risorse e riduzione dei costi ospedalieri
Limiti dichiarati	Mancanza di studi prospettici multicentrici, difficoltà di generalizzazione dei risultati, complessità computazionale elevata, barriere di implementazione nei sistemi ospedalieri esistenti
Riferimento Bibliografico	https://doi.org/10.1007/s12351-024-00884-z
Stratificazione per setting	CHIRURGICO
Inclusione studio nella Revisione	SI

ID STUDIO:	ST55
TITOLO:	Machine learning based forecast for the prediction of inpatient bed demand
Titolo in italiano:	Previsione della domanda di posti letto ospedalieri tramite machine learning

Autore/i	Manuel Tello, Eric S. Reich, Jason Puckey, Rebecca Maff, Andres Garcia-Arce, Biplab Sudhin Bhattacharya, Felipe Feijoo
Anno	2022
Paese	Stati Uniti
Tipo di studio	Studio osservazionale con sviluppo e validazione di modello predittivo
Contesto	Ospedaliero (gestione dei posti letto per pazienti ricoverati)
Specialità/Setting	Pianificazione della capacità ospedaliera e gestione dei flussi di ricovero
Tecnologia IA	Machine learning supervisionato
Algoritmo/Modello	Gradient Boosting Regressor (GBR) e Random Forest, confrontati con modelli statistici tradizionali di forecasting (ARIMA e Prophet)
Compito/Utilizzo	Previsione giornaliera e settimanale della domanda di posti letto per ottimizzare la capacità ospedaliera
Dataset	Dati real-world di ricoveri ospedalieri dal sistema sanitario di una rete ospedaliera statunitense; migliaia di osservazioni su più anni; set di training e test separati per validazione
Outcome principali	Accuratezza predittiva dei modelli ML rispetto ai metodi statistici tradizionali; capacità di anticipare picchi di domanda per migliorare la gestione delle risorse ospedaliere
Metri di valutazione	Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), confronto tra errori medi dei modelli ML e dei modelli statistici; analisi dei miglioramenti percentuali
Risultati chiave	I modelli di machine learning hanno superato nettamente le tecniche statistiche tradizionali, con riduzione significativa dell'errore di previsione. Il Gradient

Benefici dichiarati	Boosting ha mostrato le performance migliori, permettendo previsioni più accurate sia a breve che a medio termine. L'uso del modello ha consentito di anticipare picchi di domanda, riducendo mismatch tra capacità e necessità reali. Migliore allocazione dei posti letto, maggiore efficienza nella gestione dei flussi di ricovero, riduzione dei colli di bottiglia organizzativi, supporto decisionale per i manager ospedalieri
Limiti dichiarati	Studio monocentrico, necessità di validazioni multicentriche, dipendenza da dati storici che potrebbero non riflettere variazioni improvvise (es. pandemie), assenza di valutazione prospettica dell'impatto sui costi e sugli esiti clinici
Riferimento Bibliografico	https://doi.org/10.1186/s12911-022-01787-9
Stratificazione per setting	BED MANAGEMENT
Inclusione studio nella Revisione	NO

ID STUDIO:	ST56
TITOLO:	Decision analysis framework for predicting no-shows to appointments using machine learning algorithms
Titolo in italiano:	Framework di analisi decisionale per la previsione delle mancate presentazioni agli appuntamenti tramite algoritmi di machine learning

Autore/i	Carolina Deina, Flavio S. Fogliatto, Giovani J. C. da Silveira, Michel J. Anzanello
Anno	2024
Paese	Brasile e Canada (collaborazione internazionale)
Tipo di studio	Studio osservazionale retrospettivo con sviluppo e validazione di modelli ML
Contesto	Ambulatoriale
Specialità/Setting	Gestione delle prenotazioni in cliniche e ospedali
Tecnologia IA	Machine learning supervisionato
Algoritmo/Modello	Random Forest, Gradient Boosting Machine (GBM), Support Vector Machine (SVM), reti neurali artificiali (ANN) e modelli ensemble; confronto tra le performance dei diversi algoritmi
Compito/Utilizzo	Previsione delle mancate presentazioni (no-shows) agli appuntamenti per supportare la gestione delle agende ambulatoriali
Dataset	Dati real-world di 700.000 appuntamenti sanitari provenienti da un ospedale brasiliano; suddivisione in training e test set per la validazione dei modelli
Outcome principali	Accuratezza predittiva dei modelli nel classificare i pazienti a rischio di mancata presentazione
Metri di valutazione	AUC, accuratezza, precision, recall, F1-score; analisi comparativa delle performance degli algoritmi
Risultati chiave	I modelli ensemble e di boosting hanno ottenuto le performance migliori (AUC superiore a 0,80), superando nettamente i metodi tradizionali. L'uso del framework decisionale ha permesso di identificare le caratteristiche più predittive dei no-shows (es. età, storia di assenze, giorno della settimana).
Benefici dichiarati	Migliore gestione delle agende ambulatoriali, riduzione dei tempi morti e delle perdite economiche, possibilità di interventi proattivi (es. reminder mirati, ricontatto dei pazienti a rischio)
Limiti dichiarati	Studio monocentrico, dipendenza dalla qualità dei dati registrati, necessità di testare la generalizzabilità in altri contesti sanitari, mancata analisi dell'impatto economico diretto delle strategie derivate
Riferimento	https://doi.org/10.1186/s12913-023-10418-6

Bibliografico
Stratificazione per setting
Inclusione studio nella Revisione

AMBULATORIALE

SI

ID STUDIO:

ST57

TITOLO:

Key challenges for delivering clinical impact with artificial intelligence

Titolo in italiano:

Le principali sfide per ottenere un impatto clinico con l'intelligenza artificiale

Autore/i

Christopher J. Kelly, Alan Karthikesalingam, Mustafa Suleyman, Greg Corrado, Dominic King

Anno

2019

Paese

Regno Unito e Stati Uniti (collaborazione internazionale, Google DeepMind e Google Health)

Tipo di studio

Commento/analisi critica pubblicata su BMC Medicine

Contesto

Ospedaliero e ambulatoriale, con focus trasversale su diversi ambiti clinici

Specialità/Setting

Applicazioni trasversali dell'IA in medicina clinica, diagnostica, gestione dati e percorsi assistenziali

Tecnologia IA

Machine learning e deep learning

Algoritmo/Modello

Discussione generale senza sviluppo di un modello; esempi di CNN per imaging, NLP per analisi di testo clinico e modelli predittivi applicati a EHR

Compito/Utilizzo

Analisi delle barriere che limitano la traduzione dell'IA dalla ricerca alla pratica clinica e individuazione delle condizioni necessarie per un impatto reale

Dataset

Non presenta dataset originale; vengono discussi esempi da letteratura e progetti di sviluppo in collaborazione con sistemi sanitari

Outcome principali

Identificazione delle aree critiche che ostacolano l'adozione clinica dell'IA: validazione esterna, integrazione nei workflow clinici, trasparenza e spiegabilità, governance dei dati, regolamentazione e aspetti etici

Metri di valutazione

Non applicabili (studio concettuale e di analisi critica)

Risultati chiave

Perché l'IA abbia un impatto reale occorre superare la distanza tra performance in laboratorio e contesto reale, garantire interoperabilità con i sistemi clinici, affrontare bias nei dati, sviluppare modelli spiegabili e costruire fiducia tra operatori sanitari e pazienti. L'articolo individua la necessità di collaborazione stretta tra sviluppatori, clinici e regolatori.

Benefici dichiarati

L'IA ha un enorme potenziale per migliorare diagnosi, prognosi, pianificazione e gestione sanitaria; se implementata correttamente può ridurre errori, ottimizzare le risorse e aumentare l'accessibilità ai servizi

Limiti dichiarati

Mancanza di evidenze cliniche prospettiche, difficoltà di standardizzazione e validazione esterna, ostacoli regolatori e culturali, rischio di hype tecnologico non supportato da impatti concreti

Riferimento

<https://doi.org/10.1186/s12916-019-1426-2>

Bibliografico

AMBITO GENERALISTA AI IN SANITA'

Stratificazione per setting

Inclusione studio nella Revisione

NO

ID STUDIO: ST58
TITOLO: Artificial intelligence in surgery
Titolo in italiano: Intelligenza artificiale in chirurgia

Autore/i	Chris Varghese, Ewen M. Harrison, Greg O'Grady, Eric J. Topol
Anno	2024
Paese	Regno Unito, Nuova Zelanda, Stati Uniti
Tipo di studio	Revisione narrativa pubblicata su Nature Medicine
Contesto	Chirurgico
Specialità/Setting	Chirurgia generale e specialistica (inclusi sottosectori come chirurgia addominale, toracica, robotica e mini-invasiva)
Tecnologia IA	Machine learning, deep learning, computer vision, natural language processing, intelligenza artificiale integrata con robotica
Algoritmo/Modello	Discussione di CNN per visione intraoperatoria, NLP per estrazione automatica da cartelle cliniche e report chirurgici, modelli predittivi ML per esiti post-operatori, sistemi di AI chirurgica autonoma e semi-autonoma; non viene sviluppato un modello originale ma presentati esempi di applicazioni cliniche
Compito/Utilizzo	Supporto decisionale, analisi intraoperatoria, automazione di procedure robotiche, previsione delle complicanze e ottimizzazione dei percorsi chirurgici
Dataset	Evidenze tratte da studi precedenti, inclusi dataset real-world di imaging intraoperatorio, EHR e registri chirurgici; casistiche variabili da centinaia a migliaia di procedure a seconda delle applicazioni riportate
Outcome principali	L'IA in chirurgia ha migliorato la precisione tecnica, la capacità predittiva sugli esiti, la sicurezza intraoperatoria e l'efficienza organizzativa. Viene mostrato come l'integrazione di AI e robotica possa trasformare il ruolo del chirurgo e del team operatorio
Metri di valutazione	Accuratezza diagnostica e predittiva, AUC, sensibilità/specificità, tassi di successo chirurgico, confronto con tecniche tradizionali
Risultati chiave	L'IA sta trasformando la chirurgia attraverso sistemi predittivi e robotici che superano le capacità umane in termini di analisi dati e precisione intraoperatoria. Le applicazioni comprendono dall'analisi video automatizzata, al supporto clinico personalizzato, fino ai primi esempi di autonomia robotica. Tuttavia, le barriere riguardano spiegabilità, regolamentazione e fiducia clinica.
Benefici dichiarati	Miglioramento degli outcome chirurgici, riduzione delle complicanze, ottimizzazione dell'uso delle risorse e formazione chirurgica avanzata tramite simulazioni AI-driven
Limiti dichiarati	Mancanza di validazioni prospettiche multicentriche, rischio di bias nei dataset, problematiche etiche e legali sull'autonomia chirurgica, difficoltà di integrazione nei sistemi clinici reali
Riferimento Bibliografico	https://doi.org/10.1038/s41591-024-02970-3
Stratificazione per setting	CHIRURGICO
Inclusione studio nella Revisione	SI

ID STUDIO: ST59
TITOLO: Impact of data on generalization of AI for surgical intelligence applications

Titolo in italiano: **L'impatto dei dati sulla generalizzazione dell'intelligenza artificiale nelle applicazioni di chirurgia intelligente**

Autore/i	Omri Bar, Daniel Neimark, Maya Zohar, Gregory D. Hager, Ross Girshick, Gerald M. Fried, Tamir Wolf, Dotan Asselmann
Anno	2020
Paese	Stati Uniti, Canada, Israele (collaborazione internazionale)
Tipo di studio	Studio metodologico sperimentale
Contesto	Chirurgico
Specialità/Setting	Chirurgia laparoscopica e applicazioni di analisi video intraoperatoria
Tecnologia IA	Deep learning e computer vision
Algoritmo/Modello	Reti neurali convoluzionali (CNN) per il riconoscimento di fasi chirurgiche e attività intraoperatorie; valutazione di modelli addestrati su dataset provenienti da istituzioni differenti
Compito/Utilizzo	Analizzare la capacità di generalizzazione dei modelli AI per la classificazione delle fasi chirurgiche e delle attività intraoperatorie, valutando l'impatto della provenienza e della diversità dei dati di training
Dataset	Diversi dataset di video laparoscopici real-world provenienti da più centri; dimensione complessiva di diverse centinaia di interventi. Confronto tra training su dataset single-center e multi-center
Outcome principali	I modelli addestrati su dataset limitati a un singolo centro hanno mostrato scarsa capacità di generalizzazione, con riduzione significativa delle performance quando applicati a dati esterni. L'integrazione di dataset multi-istituzionali ha migliorato sensibilmente accuratezza e robustezza del modello.
Metri di valutazione	Accuratezza di classificazione, F1-score, analisi comparativa delle performance intra-dataset e cross-dataset
Risultati chiave	La qualità e la diversità dei dati sono determinanti per l'affidabilità dell'IA in chirurgia. L'uso di dataset multicentrici e rappresentativi consente modelli più generalizzabili e applicabili nella pratica clinica reale. L'articolo sottolinea la necessità di creare grandi repository condivisi di dati chirurgici.
Benefici dichiarati	Miglioramento della robustezza e trasferibilità dei sistemi AI, possibilità di implementazioni cliniche più sicure ed efficaci, riduzione del rischio di bias istituzionali
Limiti dichiarati	Dimensioni ancora limitate dei dataset disponibili, mancanza di standardizzazione nei formati video e nelle annotazioni, difficoltà etiche e legali nella condivisione dei dati clinici tra istituzioni
Riferimento Bibliografico	https://doi.org/10.1038/s41598-020-79173-6
Stratificazione per setting	CHIRURGICO
Inclusione studio nella Revisione	SI

ID STUDIO: **ST60**

TITOLO: **Bias and ethical considerations in machine learning and the automation of perioperative risk assessment**

Titolo in italiano: **Pregiudizi e considerazioni etiche nel machine learning e nell'automazione della valutazione del rischio perioperatorio**

Autore/i Vikas N. O'Reilly-Shah, Katherine R. Gentry, Andrew M. Walters, Joel Zivot, Corrie T. Anderson, Patrick J. Tighe

Anno	2020
Paese	Stati Uniti
Tipo di studio	Editoriale / Analisi critica
Contesto	Chirurgico (anestesiologia e gestione perioperatoria)
Specialità/Setting	Anestesia e medicina perioperatoria
Tecnologia IA	Machine Learning e Deep Learning
Algoritmo/Modello	Non applicabile (analisi critica, non studio sperimentale)
Compito/Utilizzo	Supporto decisionale e calcolo di risk score perioperatori
Dataset	Non applicabile (discussione teorica, esempi presi da dataset come NSQIP e STS)
Outcome principali	Identificazione delle fonti di bias nei modelli di rischio, implicazioni etiche sull'uso di algoritmi in anestesia e chirurgia
Metri di valutazione	Non riportate (trattazione concettuale)
Risultati chiave	I modelli di machine learning perioperatori rischiano di propagare bias legati a razza, genere e determinanti sociali, con potenziali impatti negativi sulla qualità e l'equità delle cure. L'opacità dei modelli complessi può rendere difficile rilevare tali bias e può portare a decisioni cliniche inappropriate.
Benefici dichiarati	L'uso dell'IA potrebbe permettere una valutazione personalizzata del rischio e migliorare il supporto decisionale in sala operatoria.
Limiti dichiarati	Mancanza di trasparenza nei dati e nei processi di sviluppo dei modelli, rischio di bias nei dataset di training, difficoltà per i clinici nel valutare la validità degli algoritmi, possibilità di uso distorto da parte di sistemi sanitari o assicurazioni.
Riferimento Bibliografico	https://doi.org/10.1016/j.bja.2020.07.040
Stratificazione per setting	CHIRURGICO
Inclusione studio nella Revisione	SI

ID STUDIO:	ST61
TITOLO:	An adaptive decision support system for outpatient appointment scheduling with heterogeneous service times
Titolo in italiano:	Un sistema adattivo di supporto decisionale per la pianificazione degli appuntamenti ambulatoriali con tempi di servizio eterogenei
Autore/i	Haolin Feng, Yiwu Jia, Teng Huang, Siyi Zhou, Hongyi Chen
Anno	2024
Paese	Cina
Tipo di studio	Studio metodologico con sviluppo e validazione di sistema decisionale AI-based
Contesto	Ambulatoriale
Specialità/Setting	Pianificazione degli appuntamenti in cliniche ospedaliere
Tecnologia IA	Machine learning e ottimizzazione adattiva
Algoritmo/Modello	Decision Support System basato su reinforcement learning adattivo e tecniche di simulazione-ottimizzazione per gestire la variabilità dei tempi di servizio
Compito/Utilizzo	Ottimizzazione della pianificazione degli appuntamenti ambulatoriali, riduzione dei tempi di attesa e miglioramento dell'utilizzo delle risorse
Dataset	Dati real-world provenienti da cliniche ambulatoriali cinesi, comprendenti storici sugli appuntamenti, durata delle visite e tempi di attesa; dimensione: diverse migliaia di appuntamenti
Outcome principali	Valutazione dell'efficienza del sistema adattivo rispetto ai metodi tradizionali di scheduling, con focus su riduzione dei tempi di attesa e incremento del tasso di

Metri di valutazione	utilizzo delle risorse Tempo medio di attesa, tasso di utilizzo delle risorse, accuratezza predittiva, confronto percentuale rispetto ai metodi convenzionali
Risultati chiave	Il sistema adattivo basato su AI ha ridotto significativamente i tempi medi di attesa e migliorato l'uso delle risorse rispetto agli approcci statici. L'algoritmo ha dimostrato capacità di adattarsi dinamicamente alla variabilità dei tempi di servizio, mostrando robustezza in scenari realistici.
Benefici dichiarati	Miglioramento dell'efficienza dei processi ambulatoriali, riduzione dei disagi per i pazienti, utilizzo ottimale delle risorse mediche e infermieristiche, supporto alle decisioni dei manager sanitari
Limiti dichiarati	Necessità di ulteriori validazioni multicentriche, rischio di complessità computazionale in contesti molto ampi, dipendenza dalla qualità dei dati storici disponibili
Riferimento Bibliografico	https://doi.org/10.1038/s41598-024-77873-x
Stratificazione per setting	AMBULATORIALE
Inclusione studio nella Revisione	SI

ID STUDIO:	ST62
TITOLO:	Clinical artificial intelligence quality improvement: towards continual monitoring and updating of AI algorithms in healthcare
Titolo in italiano:	Miglioramento della qualità dell'intelligenza artificiale clinica: verso un monitoraggio e aggiornamento continuo degli algoritmi in sanità
Autore/i	Jean Feng, Rachael V. Phillips, Ivana Malenica, Andrew Bishara, Alan E. Hubbard, Leo A. Celi, Romain Pirracchio
Anno	2022
Paese	Stati Uniti e Francia (collaborazione internazionale)
Tipo di studio	Commento / Proposta metodologica
Contesto	Ospedaliero e ambulatoriale, applicazioni trasversali in sanità
Specialità/Setting	Qualità e sicurezza dell'implementazione di sistemi AI clinici
Tecnologia IA	Machine learning e deep learning
Algoritmo/Modello	Non sviluppa un singolo algoritmo, ma propone un framework di aggiornamento continuo ("continual learning") per modelli predittivi già implementati in ambito clinico
Compito/Utilizzo	Definizione di un approccio strutturato al monitoraggio della performance dei modelli AI nel tempo, con strategie di retraining per mantenere accuratezza e affidabilità
Dataset	Discussione su dataset real-world clinici (cartelle elettroniche, dati perioperatori, outcome clinici) con esempi di applicazioni su grandi sistemi ospedalieri; non presenta un dataset originale ma cita esperienze pratiche
Outcome principali	Necessità di sviluppare un ciclo di miglioramento continuo per gli algoritmi clinici, analogamente a quanto avviene per i protocolli di qualità ospedaliera. Identifica rischi legati a "dataset shift" e obsolescenza dei modelli, che possono portare a peggioramento della performance clinica se non aggiornati.
Metri di valutazione	Accuratezza predittiva, AUC, calibrazione dei modelli, monitoraggio continuo delle metriche di performance rispetto ai dati reali correnti
Risultati chiave	La gestione degli algoritmi clinici deve essere vista come un processo dinamico, con governance strutturata per garantire aggiornamento continuo e

Benefici dichiarati	minimizzazione dei bias. Propone la creazione di infrastrutture ospedaliere dedicate al "lifecycle management" dell'IA. Maggiore affidabilità dei modelli nel tempo, riduzione del rischio clinico, miglioramento della sicurezza dei pazienti, costruzione di fiducia verso l'IA da parte degli operatori sanitari
Limiti dichiarati	Mancanza di validazioni prospettiche del framework, difficoltà di implementazione nei sistemi ospedalieri reali, complessità organizzativa e necessità di nuove figure professionali dedicate alla gestione degli algoritmi
Riferimento Bibliografico	https://doi.org/10.1038/s41746-022-00611-y
Stratificazione per setting	AMBITO GENERALISTA AI IN SANITA'
Inclusione studio nella Revisione	SI

ID STUDIO:	ST63
TITOLO:	Integrated multimodal artificial intelligence framework for healthcare applications
Titolo in italiano:	Framework multimodale integrato di intelligenza artificiale per applicazioni sanitarie

Autore/i	Luis R. Soenksen, Yu Ma, Cynthia Zeng, Leonard Boussieux, Kimberly Villalobos Carballo, Liangyuan Na, Holly M. Wiberg, Michael L. Li, Ignacio Fuentes, Dimitris Bertsimas
Anno	2022
Paese	Stati Uniti (Massachusetts Institute of Technology e collaborazioni)
Tipo di studio	Studio metodologico con proposta di framework AI
Contesto	Ospedaliero e ambulatoriale, applicazioni trasversali in diversi ambiti clinici
Specialità/Setting	Medicina predittiva, diagnostica e decisione clinica basata su dati multimodali
Tecnologia IA	Machine learning, deep learning, modelli multimodali
Algoritmo/Modello	Framework "IMPACT" (Integrated Multimodal Prediction And Classification Tool) basato su reti neurali multimodali che integrano dati clinici strutturati, immagini mediche e note testuali
Compito/Utilizzo	Predizione e classificazione clinica combinando fonti eterogenee di dati per migliorare accuratezza e generalizzabilità delle decisioni AI
Dataset	Testato su dataset real-world comprendenti oltre 200.000 pazienti, con integrazione di dati EHR, imaging e testi clinici; dataset provenienti da istituzioni statunitensi
Outcome principali	Dimostrazione che l'approccio multimodale supera le performance dei modelli unimodali; aumento dell'accuratezza predittiva e miglior calibrazione dei modelli
Metri di valutazione	AUC, accuratezza, F1-score, calibrazione predittiva; confronto tra framework multimodale e modelli tradizionali
Risultati chiave	L'integrazione multimodale ha migliorato significativamente le capacità predittive rispetto a singole fonti di dati, riducendo il rischio di bias e aumentando la robustezza clinica. Il framework ha mostrato potenzialità per applicazioni trasversali, dalla diagnosi precoce alle decisioni terapeutiche.
Benefici dichiarati	Maggior accuratezza diagnostica e prognostica, miglior generalizzazione dei modelli, possibilità di supporto clinico in molteplici specialità con un'unica architettura AI
Limiti dichiarati	Complessità computazionale, necessità di dataset ampi e ben annotati, barriere legate all'interoperabilità dei dati clinici, mancanza di validazioni prospettiche

Riferimento
Bibliografico
Stratificazione per
setting
Inclusione studio
nella Revisione

multicentriche

<https://doi.org/10.1038/s41746-022-00689-4>

AMBITO GENERALISTA AI IN SANITA'

NO

ID STUDIO: ST64

TITOLO: The digital transformation of surgery

Titolo in italiano: La trasformazione digitale della chirurgia

Autore/i

Jayson S. Marwaha, Marium M. Raza, Joseph C. Kvedar

Anno

2023

Paese

Stati Uniti

Tipo di studio

Editoriale introduttivo a collezione tematica

Contesto

Chirurgico

Specialità/Setting

Innovazioni digitali applicate alla chirurgia

Tecnologia IA

Machine learning, computer vision, realtà aumentata e virtuale, sensori e wearable devices

Algoritmo/Modello

Non viene descritto un singolo modello, ma vengono discussi esempi di machine learning per la predizione degli esiti chirurgici, computer vision per l'analisi intraoperatoria e sistemi di monitoraggio digitale

Compito/Utilizzo

Supporto decisionale, predizione degli esiti chirurgici, analisi intraoperatoria, monitoraggio remoto dei pazienti

Dataset

Si fa riferimento a dataset real-world (cartelle cliniche elettroniche, dati intraoperatori, wearable) citando studi pilota e trial randomizzati; non è riportato un dataset originale nell'editoriale

Outcome principali

L'IA e le tecnologie digitali stanno dimostrando potenziale per migliorare la predizione degli esiti chirurgici, supportare decisioni intraoperatorie, monitorare da remoto il recupero postoperatorio e ridurre le complicanze.

Metri di valutazione

AUC, accuratezza predittiva, efficacia nel monitoraggio remoto (citazioni da studi precedenti), ma l'articolo non riporta valori numerici propri.

Risultati chiave

Le tecnologie emergenti – AI, realtà aumentata, wearable e monitoraggio remoto – hanno il potenziale di trasformare la chirurgia. Tuttavia, barriere come equità, sicurezza informatica, regolamentazione e scarsità di evidenze cliniche solide limitano la diffusione.

Benefici dichiarati

Miglioramento della precisione chirurgica, riduzione dei tempi e delle complicanze postoperatorie, possibilità di monitoraggio domiciliare, evoluzione verso una chirurgia più personalizzata e sicura.

Limiti dichiarati

Mancanza di evidenze robuste di impatto clinico, rischi legati al bias, alle disuguaglianze sanitarie, alla sicurezza informatica e all'adattamento regolatorio.

Riferimento

<https://doi.org/10.1038/s41746-023-00846-3>

Bibliografico

Stratificazione per

setting

Inclusione studio

nella Revisione

CHIRURGICO

SI

ID STUDIO: ST65

TITOLO: **Development, Deployment, and Implementation of a Machine Learning Surgical Case Length Prediction Model and Prospective Evaluation**

Titolo in italiano: **Sviluppo, implementazione e valutazione prospettica di un modello di machine learning per la previsione della durata degli interventi chirurgici**

Autore/i	Hamed Zaribafzadeh, Wendy L. Webster, Christopher J. Vail, Thomas Daigle, Allan D. Kirk, Peter J. Allen, Ricardo Henao, Daniel M. Buckland
Anno	2023
Paese	Stati Uniti
Tipo di studio	Studio prospettico con implementazione reale di un modello ML
Contesto	Chirurgico
Specialità/Setting	Sala operatoria (multi-specialità, setting ambulatoriale e ospedaliero)
Tecnologia IA	Machine Learning (Gradient Boosting)
Algoritmo/Modello	Gradient Boosted Decision Trees (XGBoost) con "similarity cascade" e loss function personalizzata
Compito/Utilizzo	Previsione della durata degli interventi chirurgici per ottimizzare la programmazione
Dataset	107.898 casi chirurgici del Duke University Health System (DUHS) dal 2021–2022 (80.595 training, 27.303 test). Dati real-world da EHR.
Outcome principali	Accuratezza nella previsione della durata degli interventi chirurgici; bilanciamento tra errori di sovrastima e sottostima; miglioramento nell'efficienza operativa.
Metri di valutazione	Mean Squared Logarithmic Error (MSLE), Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), % di casi previsti entro $\pm 20\%$ della durata reale.
Risultati chiave	Il modello ha previsto il 58,7% dei casi entro il 20% della durata reale, con bilanciamento tra over e under-prediction. Nella valutazione prospettica (33.815 casi), ha ridotto dell'11,2% le sottostime, aumentato del 5,9% le previsioni corrette ed è stato progressivamente adottato dai programmatori. Ha portato a minori ore di straordinario (-485 ore inpatient, -53 ore ambulatoriali) e minore errore totale medio per sala operatoria al giorno.
Benefici dichiarati	Miglior utilizzo delle sale operatorie, riduzione dei tempi di attesa di pazienti e chirurghi, risparmio di costi legati a straordinari, possibilità di aggiungere nuovi casi sfruttando i tempi risparmiati.
Limiti dichiarati	Necessità di continuo aggiornamento del modello con nuovi dati; variabilità legata a CPT, nuovi chirurghi e tecniche; affidabilità condizionata dalla qualità dei dati inseriti; studio monocentrico da validare in altri contesti.
Riferimento Bibliografico	https://doi.org/10.1097/SLA.0000000000005936
Stratificazione per setting	CHIRURGICO
Inclusione studio nella Revisione	SI

ID STUDIO: ST66

TITOLO: **Artificial Intelligence for Healthcare and Social Services: Optimizing Resources and Promoting Sustainability**

Titolo in italiano: **Intelligenza artificiale per sanità e servizi sociali: ottimizzazione delle risorse e promozione della sostenibilità**

Autore/i	Eliseo Sciarretta, Riccardo Mancini, Emilio Greco
Anno	2022
Paese	Italia
Tipo di studio	Revisione sistematica della letteratura
Contesto	Sanitario e sociale
Specialità/Setting	Applicazioni trasversali in prevenzione, diagnosi, trattamento, monitoraggio e management sanitario
Tecnologia IA	Machine Learning, Natural Language Processing, robotica, IoT e sistemi ibridi
Algoritmo/Modello	Non viene descritto un modello unico; la revisione affronta varie tipologie (ML supervisionato e non supervisionato, NLP, robotica assistita e autonoma)
Compito/Utilizzo	Prevenzione, supporto decisionale clinico, personalizzazione delle terapie, monitoraggio remoto tramite wearable, ottimizzazione della gestione delle risorse
Dataset	Revisione di 40 studi pubblicati negli ultimi 10 anni; fonti da PubMed, SAGE, Springer. Dati prevalentemente real-world (big data clinici, EHR, wearable)
Outcome principali	L'IA migliora la prevenzione (es. rischio suicidario), riduce errori diagnostici, personalizza i trattamenti, consente monitoraggio remoto dei pazienti cronici e ottimizza processi amministrativi.
Metri di valutazione	Non riportate metriche specifiche (sintesi di studi secondari)
Risultati chiave	L'IA può rivoluzionare i servizi sanitari, aumentandone sostenibilità ed efficienza. Ha già mostrato risultati promettenti in prevenzione, diagnosi e gestione clinica, ma presenta criticità legate a maturità tecnologica, fiducia degli operatori e carenze regolatorie.
Benefici dichiarati	Miglioramento della qualità e continuità assistenziale, riduzione dei costi, maggiore sostenibilità del sistema sanitario e sociale, personalizzazione delle cure
Limiti dichiarati	Algoritmi ancora immaturi, necessità di grandi dataset, problemi di privacy, mancanza di regolamentazione, rischio di bias e disuguaglianze sanitarie, fiducia limitata di operatori e pazienti
Riferimento Bibliografico	https://doi.org/10.3390/su142416464
Stratificazione per setting	AMBITO GENERALISTA AI IN SANITA'
Inclusione studio nella Revisione	NO

ID STUDIO:	ST67
TITOLO:	Predicting non-attendance in hospital outpatient appointments using deep learning approach
Titolo in italiano:	Predizione delle mancate presenze agli appuntamenti ospedalieri ambulatoriali tramite approccio di deep learning

Autore/i	M. Dashtban, Weizi Li
Anno	2022
Paese	Regno Unito
Tipo di studio	Studio sperimentale con sviluppo e validazione di modello predittivo
Contesto	Contesto: Ambulatoriale
Specialità/Setting	Ambulatoriale
Tecnologia IA	Pianificazione degli appuntamenti ospedalieri Deep Learning

Algoritmo/Modello	Reti neurali profonde (Deep Neural Networks) confrontate con modelli di machine learning tradizionali (Random Forest, Logistic Regression, SVM)
Compito/Utilizzo	Previsione delle mancate presenze (no-show) agli appuntamenti ambulatoriali per migliorare la pianificazione e ridurre sprechi di risorse
Dataset	Dataset reale proveniente da un ospedale del Regno Unito, comprendente oltre 1,2 milioni di appuntamenti ambulatoriali dal 2015 al 2019, con variabili cliniche, demografiche e storiche di comportamento dei pazienti
Outcome principali	Identificazione dei pazienti a rischio di mancata presentazione, miglioramento dell'efficienza del sistema tramite riduzione dei no-show
Metri di valutazione	AUC, accuratezza, precisione, recall, F1-score; il modello di deep learning ha superato le performance degli algoritmi tradizionali
Risultati chiave	Il modello di deep learning ha ottenuto migliori risultati predittivi rispetto ai metodi convenzionali, dimostrando la capacità di gestire grandi dataset complessi e identificare pattern non lineari nel comportamento dei pazienti.
Benefici dichiarati	Riduzione dei costi dovuti a mancata utilizzazione delle risorse, maggiore efficienza nella pianificazione degli appuntamenti, possibilità di interventi mirati sui pazienti ad alto rischio di no-show
Limiti dichiarati	Necessità di validazioni multicentriche, rischio di overfitting, dipendenza dalla qualità e completezza dei dati storici, potenziale mancanza di trasparenza interpretativa dei modelli di deep learning
Riferimento Bibliografico	https://doi.org/10.1080/20476965.2021.1924085
Stratificazione per setting	AMBULATORIALE
Inclusione studio nella Revisione	SI

ID STUDIO:	ST68
TITOLO:	Exploring the Impact of Artificial Intelligence on Global Health and Enhancing Healthcare in Developing Nations
Titolo in italiano:	Esplorare l'impatto dell'intelligenza artificiale sulla salute globale e migliorare l'assistenza sanitaria nei paesi in via di sviluppo
Autore/i	Varisha Zuhair, Areesha Babar, Rabbiya Ali, Malik Olatunde Oduoye, Zainab Noor, Kitumaini Chris, Inibehe Ime Okon, Latif Ur Rehman
Anno	2024
Paese	Pakistan, Nigeria, RDC, Camerun (collaborazione internazionale)
Tipo di studio	Revisione narrativa
Contesto	Sanitario globale, con focus su sistemi a basse risorse
Specialità/Setting	Applicazioni trasversali in diagnostica, monitoraggio remoto, telemedicina e gestione delle risorse
Tecnologia IA	Machine Learning, Deep Learning, Natural Language Processing, Computer Vision
Algoritmo/Modello	Non descritto un singolo modello; vengono riportati esempi di CNN per diagnosi radiologica, NLP per telemedicina multilingua, e sistemi predittivi per malattie infettive
Compito/Utilizzo	Supporto decisionale clinico, diagnosi precoce, ottimizzazione risorse, miglioramento dell'accesso alle cure
Dataset	Riferimenti a dataset globali e real-world, come immagini radiologiche (es. tubercolosi), dati epidemiologici su malattie infettive, e cartelle elettroniche parzialmente digitalizzate; mancano dataset originali propri
Outcome principali	L'IA ha il potenziale di migliorare la diagnosi precoce di patologie (TB, malaria,

Metri di valutazione	HIV), ridurre la mortalità materno-infantile, ottimizzare la distribuzione delle risorse e supportare la telemedicina in contesti a bassa connettività.
Risultati chiave	Non riportate metriche numeriche; la revisione cita risultati di studi precedenti (AUC >0,9 per TB, accuratezza >90% per malaria in test locali)
Benefici dichiarati	L'IA può contribuire a colmare le disparità sanitarie globali, ma richiede infrastrutture tecnologiche, interoperabilità dei dati e formazione locale. L'articolo evidenzia anche rischi legati a bias culturali e mancanza di rappresentatività dei dataset globali.
Limiti dichiarati	Accesso più equo alle cure, diagnosi precoce in aree remote, rafforzamento dei sistemi sanitari fragili, riduzione delle disuguaglianze sanitarie globali
Riferimento Bibliografico	Scarsa disponibilità di dataset locali, infrastrutture digitali limitate, dipendenza da tecnologie sviluppate nei paesi ad alto reddito, rischi di colonialismo digitale
Stratificazione per setting	https://doi.org/10.1177/21501319241245847
Inclusione studio nella Revisione	AMBITO GENERALISTA AI IN SANITA' NO

ID STUDIO:	ST69
TITOLO:	Exploring the Impact of Artificial Intelligence on Global Health and Enhancing Healthcare in Developing Nations
Titolo in italiano:	Esplorare l'impatto dell'intelligenza artificiale sulla salute globale e migliorare l'assistenza sanitaria nei paesi in via di sviluppo

Autore/i	Fatema Mustansir Dawoodbhoj, Jack Delaney, Paulina Cecula, Jiakun Yu, Iain Peacock, Joseph Tan, Benita Cox
Anno	2021
Paese	Regno Unito
Tipo di studio	Revisione narrativa e analisi applicativa
Contesto	Ospedaliero, unità psichiatriche per pazienti acuti
Specialità/Setting	Salute mentale, gestione dei percorsi assistenziali e organizzazione dei reparti
Tecnologia IA	Machine Learning, Deep Learning, Natural Language Processing, modelli predittivi
Algoritmo/Modello	Non un singolo modello, ma discussione di diverse applicazioni: reti neurali, NLP per analisi di testi clinici, modelli predittivi di rischio per ottimizzare i flussi
Compito/Utilizzo	Previsione della domanda di posti letto, riduzione dei tempi di attesa, ottimizzazione del flusso tra ricovero, trattamento e dimissione
Dataset	Dati real-world provenienti dal NHS (cartelle cliniche elettroniche, dati di ricovero psichiatrico e amministrativi). Non viene riportata una dimensione unica ma sono citati esempi da dataset locali di grandi ospedali del NHS
Outcome principali	Miglioramento della gestione dei flussi dei pazienti, riduzione dei tempi di attesa per il ricovero, ottimizzazione delle dimissioni, supporto ai manager per la programmazione delle risorse
Metri di valutazione	Accuratezza predittiva, AUC, tempi medi di attesa, tasso di occupazione dei posti letto; riportati valori da studi secondari inclusi nella revisione
Risultati chiave	L'IA ha mostrato capacità di prevedere la domanda futura di letti e le tempistiche di dimissione, migliorando la fluidità dei percorsi nei reparti psichiatrici. Gli algoritmi NLP si sono rivelati utili per estrarre informazioni cliniche rilevanti dai testi delle cartelle.
Benefici dichiarati	Aumento dell'efficienza organizzativa, riduzione della pressione su operatori e

Limiti dichiarati	strutture, miglioramento della continuità assistenziale per i pazienti con patologie psichiatriche acute Mancanza di validazioni multicentriche, scarsa interoperabilità dei sistemi informativi del NHS, potenziale resistenza degli operatori all'adozione di strumenti predittivi, rischio di bias nei modelli
Riferimento Bibliografico	https://doi.org/10.1016/j.heliyon.2021.e06993
Stratificazione per setting	AMBITO GENERALISTA AI IN SANITA'
Inclusione studio nella Revisione	NO

ID STUDIO:	ST70
TITOLO:	Artificial Intelligence in Outpatient Primary Care: A Scoping Review on Applications, Challenges, and Future Directions
Titolo in italiano:	Intelligenza artificiale nelle cure primarie ambulatoriali: una scoping review su applicazioni, sfide e prospettive future

Autore/i	Kevin B. Johnson
Anno	2025
Paese	Stati Uniti
Tipo di studio	Scoping Review (Preprint)
Contesto	Ambulatoriale, cure primarie
Specialità/Setting	Medicina di base, medicina generale, gestione pazienti cronici
Tecnologia IA	Machine Learning, Deep Learning, Natural Language Processing, sistemi predittivi e di supporto decisionale
Algoritmo/Modello	Non un singolo modello, ma ampia rassegna: algoritmi predittivi per rischio clinico, NLP per estrazione di informazioni da note cliniche, modelli di stratificazione dei pazienti
Compito/Utilizzo	Supporto decisionale clinico, gestione cronici, stratificazione del rischio, ottimizzazione dei flussi di lavoro, miglioramento della comunicazione medico-paziente
Dataset	Revisione di studi basati su dataset real-world provenienti da EHR, registri sanitari nazionali, e sistemi di sorveglianza epidemiologica; dimensioni variabili (da migliaia a milioni di pazienti)
Outcome principali	L'IA può migliorare la diagnosi precoce, la gestione proattiva dei pazienti cronici, la prevenzione, la programmazione delle visite e la comunicazione tra operatori e pazienti.
Metri di valutazione	Citazione di AUC, accuratezza, precisione, recall, ma senza valori specifici propri (dipendenti dagli studi inclusi)
Risultati chiave	L'IA ha dimostrato di poter supportare le cure primarie attraverso predizioni affidabili, automazione di compiti amministrativi e analisi predittiva della popolazione assistita. Tuttavia, permangono sfide significative in termini di implementazione reale, interoperabilità e fiducia da parte dei professionisti.
Benefici dichiarati	Miglioramento della qualità delle cure, riduzione dei tempi di attesa, personalizzazione dell'assistenza, maggiore efficienza nei flussi di lavoro
Limiti dichiarati	Mancanza di validazioni cliniche robuste, rischio di bias nei dataset, difficoltà di integrazione nei sistemi sanitari reali, questioni etiche e di trasparenza
Riferimento Bibliografico	https://doi.org/10.1101/2025.05.12.25327223
Stratificazione per	CURE PRIMARIE

setting Inclusione studio nella Revisione	NO
ID STUDIO:	ST71
TITOLO:	Predicting patient outcomes and risk for revision surgery after hip and knee replacement surgery: study protocol for a comparison of modelling approaches using the Swiss National Joint Registry (SIRIS)
Titolo in italiano:	Predizione degli esiti e del rischio di revisione dopo protesi d'anca e di ginocchio: protocollo di studio per un confronto tra approcci di modellizzazione basati sul registro nazionale svizzero SIRIS
Autore/i	Léonie Hofstetter, Nathalie Schweyckart, Christof Seiler, Christian Brand, Laura C. Rosella, Mazda Farshad, Milo A. Puhan, Cesar A. Hincapié
Anno	2025
Paese	Svizzera
Tipo di studio	Protocollo di studio (comparazione tra modelli predittivi)
Contesto	Chirurgico
Specialità/Setting	Ortopedia, chirurgia protesica (anca e ginocchio)
Tecnologia IA	Machine Learning
Algoritmo/Modello	Random Forest, Extreme Gradient Boosting (XGBoost), confrontati con regressione logistica
Compito/Utilizzo	Predizione degli esiti riferiti dai pazienti (PROMs) e del rischio di revisione chirurgica dopo artroplastica totale
Dataset	Swiss National Joint Registry (SIRIS), 2015–2023, con copertura del 98% degli interventi a livello nazionale; dati real-world di migliaia di pazienti sottoposti a THA e TKA
Outcome principali	Identificazione di fattori predittivi di dolore persistente, insoddisfazione post-operatoria e rischio di revisione precoce; valutazione della capacità dei modelli ML di migliorare l'accuratezza rispetto alla regressione logistica
Metri di valutazione	Accuratezza predittiva, calibrazione, curva ROC/AUC, decision curve analysis
Risultati chiave	Lo studio è in fase di protocollo, pertanto non presenta risultati definitivi; si propone di chiarire se ML apporta benefici concreti rispetto ai metodi statistici tradizionali nella predizione di esiti e revisioni protesiche
Benefici dichiarati	Potenziare il miglioramento della personalizzazione delle cure, supporto al decision making condiviso, ottimizzazione delle risorse e riduzione dei reinterventi non necessari
Limiti dichiarati	Mancanza di evidenze empiriche al momento (protocollo), rischio di bias legato alla qualità dei dati di registro, necessità di validazione esterna dei modelli
Riferimento Bibliografico	https://doi.org/10.1186/s41512-025-00200-z
Stratificazione per setting	CHIRURGICO
Inclusione studio nella Revisione	SI

ID STUDIO: ST72
TITOLO: The role of artificial intelligence in enhancing clinical nursing care: A scoping review

Titolo in italiano: Il ruolo dell'intelligenza artificiale nel migliorare l'assistenza infermieristica clinica: una scoping review

Autore/i	Zi Qi Pamela Ng, Li Ying Janice Ling, Han Shi Jocelyn Chew, Ying Lau
Anno	2021
Paese	Singapore
Tipo di studio	Scoping Review
Contesto	Ospedaliero e ambulatoriale
Specialità/Setting	Assistenza infermieristica clinica in diversi ambiti (monitoraggio pazienti, gestione cronici, supporto decisionale)
Tecnologia IA	Machine Learning, Deep Learning, Natural Language Processing, sistemi predittivi
Algoritmo/Modello	Non un singolo modello; revisione di diversi approcci, tra cui reti neurali, NLP per analisi testi clinici, modelli predittivi di rischio
Compito/Utilizzo	Supporto decisionale clinico, monitoraggio dei pazienti, ottimizzazione dei carichi di lavoro, predizione di eventi avversi
Dataset	Studi inclusi basati su dati real-world da cartelle cliniche elettroniche, sensori indossabili e dati amministrativi sanitari; dimensioni variabili tra migliaia e milioni di record
Outcome principali	L'IA migliora la capacità predittiva degli infermieri nella gestione del rischio, consente un monitoraggio più accurato dei pazienti e riduce il carico di lavoro routinario tramite automazione
Metri di valutazione	Variano negli studi analizzati: AUC, accuratezza, sensibilità, specificità, performance predittiva
Risultati chiave	L'IA ha dimostrato benefici nel supportare l'assistenza infermieristica, migliorando efficienza e sicurezza. Tuttavia, la traduzione in pratica clinica rimane limitata da mancanza di fiducia, scarsa formazione e barriere organizzative
Benefici dichiarati	Miglioramento della qualità dell'assistenza, riduzione degli errori clinici, alleggerimento del carico amministrativo per gli infermieri, maggiore tempestività negli interventi
Limiti dichiarati	Scarsa validazione clinica, difficoltà di implementazione su larga scala, resistenza del personale, questioni etiche e di trasparenza degli algoritmi
Riferimento Bibliografico	https://doi.org/10.1111/jonm.13425
Stratificazione per setting	INFERMIERISTICO
Inclusione studio nella Revisione	NO

ID STUDIO: ST73
TITOLO: Will artificial intelligence solve the human resource crisis in healthcare?
Titolo in italiano: L'intelligenza artificiale risolverà la crisi delle risorse umane in sanità?

Autore/i	Bertalan Meskó, Gergely Hetényi, Zsuzsanna Gyórfy
Anno	2018
Paese	Ungheria

Tipo di studio	Articolo di dibattito (essay / debate)
Contesto	Trasversale (ospedaliero, territoriale, sanitario in generale)
Specialità/Setting	Sanità pubblica, workforce management
Tecnologia IA	Artificial Narrow Intelligence (ANI), Machine Learning, Deep Learning, assistenti cognitivi
Algoritmo/Modello	Non uno solo; esempi includono reti neurali profonde per diagnostica e analisi immagini, IBM Watson, DeepMind Health, Atomwise
Compito/Utilizzo	Supporto decisionale, diagnosi, gestione big data, ottimizzazione delle risorse umane e processi sanitari
Dataset	Esempi citati da progetti reali (scansioni oculari Moorfields, dati assicurativi nei Paesi Bassi, database genetici, trial farmacologici con AI); fonti real-world, non simulazioni
Outcome principali	L'IA può ridurre costi, velocizzare diagnosi, ottimizzare percorsi assistenziali e sostenere la formazione, ma non sostituirà la forza lavoro umana
Metri di valutazione	Non applicabili in quanto studio di dibattito; discussione basata su esempi di applicazioni reali già in uso
Risultati chiave	L'IA non elimina il fabbisogno di professionisti sanitari ma può ridurre il carico attraverso automazione, miglior diagnosi e supporto ai processi decisionali. È probabile che chi utilizza l'IA sostituirà chi non la usa, più che l'IA sostituire direttamente i professionisti.
Benefici dichiarati	Potenziale riduzione dei costi, miglioramento dell'efficienza, maggiore accesso a cure sicure e tempestive, sostegno al personale sanitario contro burnout e carichi eccessivi
Limiti dichiarati	Mancanza di prove solide, rischi etici e legali, costi elevati per paesi a basso reddito, impossibilità di sostituire empatia e rapporto umano, mancanza di un quadro regolatorio chiaro
Riferimento Bibliografico	https://doi.org/10.1186/s12913-018-3359-4
Stratificazione per setting	GESTIONE RISORSE UMANE GENERALISTA
Inclusione studio nella Revisione	NO

ID STUDIO:	ST74
TITOLO:	Managing outpatient flow via an artificial intelligence enabled solution
Titolo in italiano:	Gestione dei flussi ambulatoriali tramite una soluzione basata su intelligenza artificiale artificiale risolverà la crisi delle risorse umane in sanità?
Autore/i	Ling Li, Fatou Diouf, Anjee Gorkhali
Anno	2022
Paese	Stati Uniti
Tipo di studio	Studio applicativo con sviluppo di soluzione AI e validazione sul campo
Contesto	Ambulatoriale
Specialità/Setting	Gestione dei flussi di pazienti in cliniche e ospedali
Tecnologia IA	Machine Learning, Deep Learning, modelli predittivi e ottimizzazione
Algoritmo/Modello	Random Forest, Support Vector Machine, Reti neurali profonde; integrazione in un sistema AI-enabled per la gestione dei flussi
Compito/Utilizzo	Previsione dei tempi di attesa, ottimizzazione dell'allocazione delle risorse, miglioramento della programmazione delle visite

Dataset	Dati real-world provenienti da cliniche ambulatoriali (dataset multiclinico, oltre 100.000 appuntamenti con informazioni su tempi di visita, caratteristiche pazienti e personale medico)
Outcome principali	Riduzione dei tempi medi di attesa, aumento della puntualità delle visite, miglior utilizzo delle risorse umane e strutturali
Metri di valutazione	Accuratezza predittiva dei modelli, tempi medi di attesa pre- e post-implementazione, riduzione delle deviazioni dagli orari programmati
Risultati chiave	L'adozione della soluzione AI ha migliorato significativamente la gestione dei flussi ambulatoriali, consentendo una riduzione consistente dei tempi di attesa e una maggiore efficienza operativa rispetto ai sistemi tradizionali
Benefici dichiarati	Miglioramento dell'esperienza del paziente, ottimizzazione delle risorse disponibili, riduzione dei costi associati a inefficienze organizzative
Limiti dichiarati	Necessità di ulteriori validazioni multicontesto, rischio di sovra-dipendenza dagli algoritmi, necessità di continuo aggiornamento dei modelli con dati recenti
Riferimento Bibliografico	https://doi.org/10.1002/sres.2870
Stratificazione per setting	AMBULATORIALE
Inclusione studio nella Revisione	SI